IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of:	)	
	)	Group Art Unit: 1652
Alan BERRY et al.	)	
	)	Examiner: Christian L. Fronda
Application No.: 09/341,600	)	
	)	Confirmation No.: 5327
Filed: September 15, 1999	)	
	)	
Attorney Docket No.: 3161-18-PUS	)	APPEAL BRIEF
	)	
For: PROCESS FOR PRODUCTION OF N-GLUCOSAMINE	)	***Submitted Via Electronic Filing***

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

Dear Sir:

This Appeal Brief is filed in furtherance of the Notice of Appeal filed in the above-identified application on August 29, 2007. A Notice of Panel Decision from Pre-Appeal Brief Review instructing Appellant to proceed to the Board of Patent Appeals and Interferences was mailed on September 26, 2007. Therefore, the time period for filing this Appeal Brief, after being extended five months by the extension of time filed herewith, expires March 31, 2008 (March 29 and 30, 2008, being a Saturday and Sunday, respectively).

The required fees in the amount of $510.00 for filing a brief in support of an appeal under 37 C.F.R. § 41.20(b)(2) and $2230.00 for a five-month extension of time are being submitted herewith via EFS-Web. No additional fees are believed to be due in connection with this filing, but if fees are due, please debit Deposit Account No. 19-1970.

The Table of Contents begins on page 2 of this paper and defines the sections of this paper.

TABLE OF CONTENTS

I. REAL PARTY IN INTEREST

The real party in interest is **Arkion Life Sciences LLC d/b/a Bio-Technical Resources Division**, the assignee of record, the Assignment being recorded with the United States Patent Office at Reel 012928, Frame 0887.

II. RELATED APPEALS AND INTERFERENCES

None.

III. STATUS OF CLAIMS

The status of the claims in the application is:

A. TOTAL NUMBER OF CLAIMS IN THE APPLICATION

Claims in the application are: 1-91

B. CURRENT STATUS OF THE CLAIMS:

1.	Claims canceled:	1-39 and 73
2.	Claims withdrawn:	None
3.	Claims pending:	40-72 and 74-91
4.	Claims allowed:	None
5.	Claims rejected:	40-72 and 74-91

C. CLAIMS ON APPEAL

The claims on appeal are: 40-72 and 74-91

IV. STATUS OF AMENDMENTS

No claim amendments have been filed subsequent to the final rejection mailed on May 2, 2007.

V. SUMMARY OF CLAIMED SUBJECT MATTER

Concise explanations of the subject matter recited in each of the pending independent claims (claims 40, 67, 70, 71, 82 and 87) are presented below. Concise explanations are also provided for dependent claims 53, 54 and 72, for which separate arguments for patentability are being presented herein.

Claim 40

Claim 40 is directed to a method to produce glucosamine by fermentation (Specification: page 8, lines 1-12).

The method comprises a step of culturing a bacterium or yeast in a fermentation medium comprising assimilable sources of carbon, nitrogen and phosphate (Specification: page 14, line 32, to page 15, line 11; page 25, lines 1-21).

The bacterium or yeast comprises at least one genetic modification that increases the activity of glucosamine-6-phosphate synthase compared to the unmodified glucosamine-6-phosphate synthase (Specification: page 15, line 21, to page 18, line 20).

The genetic modification is either of the following:

(1) the transformation of the microorganism with a recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase, wherein the synthase has glucosamine-6-phosphate synthase activity (Specification: Example 2, pages 48 to 55; page 15, line 21, to page 17, line 7); or

(2) the genetic modification of the coding region of a gene encoding a bacterial or yeast glucosamine-6-phosphate synthase (Specification: Examples 5 and 6, pages 59 to 66; page 3, lines 4-13; page 17, line 8, to page 18, line 20; page 38, lines 9-27).

The genetic modification of the coding region results in at least one nucleic acid modification selected from the group consisting of deletion, insertion, and substitution of at least one nucleotide of said coding region of the gene (Specification: page 10, line 20, to page 11, line 14).

The at least one nucleotide modification results in increased glucosamine-6-phosphate synthase activity compared to the unmodified glucosamine-6-phosphate synthase (Specification: Examples 5 and 6, pages 59 to 66; page 15, line 21, to page 18, line 20).

The step of culturing produces and accumulates a product selected from the group consisting of glucosamine-6-phosphate and glucosamine from the bacterium or yeast (Specification: Examples 3 and 4, pages 55 to 59; page 9, lines 1-20; page 24, lines 22 to 34; page 31, lines 5-15).

The method also comprises a step of recovering and purifying the product (Specification: Page 28, line 25, to page 30, line 2).

Claim 53

Claim 53 is directed to the method to produce glucosamine by fermentation as recited in claim 40, but further specifies that the genetic modification comprises the transformation of the bacterium or yeast with a recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase that has glucosamine-6-

phosphate synthase enzymatic activity (Specification: Example 2, pages 48 to 55; page 15, line 21, to page 17, line 7).

The recombinant nucleic acid molecule is operatively linked to a transcription control sequence (Specification: page 39, line 32, to page 40, line 15).

Claim 54

Claim 54 is directed to the method to produce glucosamine by fermentation as recited in claim 53, but further specifies that the recombinant nucleic acid molecule is integrated into the genome of said bacterium or yeast (Specification: page 34, line 10, to page 35, line 26).

Claim 67

Claim 67 is directed to a recombinant bacterium or yeast for producing glucosamine by a biosynthetic process (Specification: page 31, line 6, to page 32, line 6).

The bacterium or yeast is transformed with a recombinant nucleic acid molecule comprising a nucleic acid sequence encoding a bacterial or yeast glucosamine-6-phosphate synthase (Specification: page 31, line 6, to page 32, line 6).

The nucleic acid sequence being operatively linked to a transcription control sequence (Specification: page 34, lines 10 to 35).

The nucleic acid sequence comprises, in the coding region of the nucleic acid sequence, a genetic modification (Specification: Examples 5 and 6, pages 59 to 66; page 3, lines 4-13; page 17, line 8, to page 18, line 20; page 38, lines 9-27).

The genetic modification increases the activity of said glucosamine-6-phosphate synthase compared to the unmodified glucosamine-6-phosphate synthase (Specification: Examples 5 and 6, pages 59 to 66; page 15, line 21, to page 18, line 20).

The expression of said nucleic acid sequence increases production of glucosamine by said bacterium or yeast (Specification: Page 10, line 7, to page 11, line 14).

Claim 70

Claim 70 is directed to a method to produce glucosamine by fermentation (Specification: page 8, lines 1-12).

The method comprises a step of culturing a bacterium or yeast in a fermentation medium comprising assimilable sources of carbon, nitrogen and phosphate (Specification: page 14, line 32, to page 15, line 11; page 25, lines 1-21).

The bacterium or yeast comprises at least one genetic modification that increases the activity of glucosamine-6-phosphate synthase compared to the unmodified glucosamine-6-phosphate synthase (Specification: page 15, line 21, to page 18, line 20).

The genetic modification is either of the following:

(1) the transformation of the microorganism with a recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase, wherein the synthase has glucosamine-6-phosphate synthase activity (Specification: Example 2, pages 48 to 55; page 15, line 21, to page 17, line 7); or

(2) the genetic modification of a gene encoding a bacterial or yeast glucosamine-6-phosphate synthase that increases the activity of said glucosamine-6-phosphate

synthase compared to the unmodified glucosamine-6-phosphate synthase (Specification: Examples 5 and 6, pages 59 to 66; page 3, lines 4-13; page 17, line 8, to page 18, line 20).

The genetically modified bacterium or yeast of (2) is produced by the steps of:

(A) generating modifications in an isolated nucleic acid molecule comprising a nucleic acid sequence encoding a bacterial or yeast glucosamine-6-phosphate synthase to create a plurality of modified nucleic acid sequences (Specification: Example 5, pages 60 to 61; page 36, line 17, to page 38, line 16);

(B) transforming bacteria or yeast with the modified nucleic acid sequences to produce genetically modified bacteria or yeast (Specification: Examples 5 and 6, pages 59 to 66; page 38, line 33, to page 39, line 20);

(C) screening the genetically modified bacteria or yeast for glucosamine-6-phosphate synthase activity (Specification: Example 2, pages 53 to 55; page 18, lines 10-16); and

(D) selecting the genetically modified bacteria or yeast that have increased glucosamine-6-phosphate synthase activity compared to the unmodified glucosamine-6-phosphate synthase (Specification: Example 2, page 55; Example 6, pages 64-66; page 33, lines 3-21).

The step of culturing produces a product selected from the group consisting of glucosamine-6-phosphate and glucosamine from the bacterium or yeast (Specification: Examples 3 and 4, pages 55 to 59; page 9, lines 1-20; page 24, lines 22 to 34; page 31, lines 5-15).

The method also comprises a step of recovering the product (Specification: Page 28, line 25, to page 30, line 2).

Claim 71

Claim 71 is directed to a method to produce glucosamine by fermentation (Specification: page 8, lines 1-12).

The method comprises a step of culturing a bacterium or yeast in a fermentation medium comprising assimilable sources of carbon, nitrogen and phosphate (Specification: page 14, line 32, to page 15, line 11; page 25, lines 1-21).

The bacterium or yeast that expresses a recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase, wherein said synthase has glucosamine-6-phosphate synthase enzymatic activity (Specification: Example 2, pages 48 to 55; page 15, line 21, to page 17, line 7).

The step of culturing produces and accumulates a product selected from the group consisting of glucosamine-6-phosphate and glucosamine from the bacterium or yeast (Specification: Examples 3 and 4, pages 55 to 59; page 9, lines 1-20; page 24, lines 22 to 34; page 31, lines 5-15).

The method also comprises a step of recovering and purifying the product (Specification: Page 28, line 25, to page 30, line 2).

Claim 72

Claim 72 is directed to the method to produce glucosamine by fermentation as recited in claim 71, but further specifies that the glucosamine-6-phosphate synthase comprises a genetic modification that reduces glucosamine-6-phosphate product inhibition of the glucosamine-6-phosphate synthase compared to the unmodified

glucosamine-6-phosphate synthase (Specification: Examples 5 and 6, pages 59 to 66; page 17, line 25, to page 18, line 20).

Claim 82

Claim 82 is directed to a method to produce glucosamine by fermentation (Specification: page 8, lines 1-12).

The method comprises a step of culturing a bacterium or yeast in a fermentation medium comprising assimilable sources of carbon, nitrogen and phosphate (Specification: page 14, line 32, to page 15, line 11; page 25, lines 1-21).

The bacterium or yeast that expresses a recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase, wherein said synthase has glucosamine-6-phosphate synthase enzymatic activity (Specification: Example 2, pages 48 to 55; page 15, line 21, to page 17, line 7).

The glucosamine-6-phosphate synthase comprises a genetic modification that reduces the glucosamine-6-phosphate product inhibition of the glucosamine-6-phosphate synthase compared to the unmodified glucosamine-6-phosphate synthase (Specification: Examples 5 and 6, pages 59 to 66; page 17, line 25, to page 18, line 20).

The step of culturing produces and accumulates a product selected from the group consisting of glucosamine-6-phosphate and glucosamine from the bacterium or yeast (Specification: Examples 3 and 4, pages 55 to 59; page 9, lines 1-20; page 24, lines 22 to 34; page 31, lines 5-15).

The method also comprises a step of recovering and purifying the product (Specification: Page 28, line 25, to page 30, line 2).

Claim 87

Claim 87 is directed to a method to produce glucosamine by fermentation (Specification: page 8, lines 1-12).

The method comprises a step of culturing a bacterium or yeast in a fermentation medium comprising assimilable sources of carbon, nitrogen and phosphate (Specification: page 14, line 32, to page 15, line 11; page 25, lines 1-21).

The bacterium or yeast that expresses a recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase, wherein said synthase has glucosamine-6-phosphate synthase enzymatic activity (Specification: Example 2, pages 48 to 55; page 15, line 21, to page 17, line 7).

The step of culturing produces and accumulates a product selected from the group consisting of glucosamine-6-phosphate and glucosamine from the bacterium or yeast (Specification: Examples 3 and 4, pages 55 to 59; page 9, lines 1-20; page 24, lines 22 to 34; page 31, lines 5-15).

The method also comprises a step of recovering and purifying the product (Specification: Page 28, line 25, to page 30, line 2).

VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL

The two issues on appeal are as follows:

1. Whether claims 40-70, 72, 74-77 and 80-91 are unpatentable under 35 U.S.C. § 112, first paragraph, as failing to comply with the enablement requirement.

2. Whether claims 40, 53, 55, 62-64, 71, 75-79 and 87-89 are unpatentable under 35 U.S.C. § 102(b) as anticipated by Dutka-Malen et al.

VII. ARGUMENT

A. The Specification Provides Enabling Support for the Full Scope of Claims 40-70, 72, 74-77 and 80-91

Claims 40-70, 72, 74-77 and 80-91 were rejected under 35 U.S.C. § 112, first paragraph, as failing to comply with the enablement requirement. According to the Office, the specification does not provide guidance regarding the specific type of genetic modification to perform on the specific codons within the coding region of a polynucleotide encoding glucosamine-6-phosphate synthase. Office Action dated May 2, 2007, at 3-4.

To satisfy the enablement requirement, the specification must contain sufficient disclosure to enable one skilled in the art to make and use the claimed invention without undue experimentation. M.P.E.P. § 2164. A determination of whether the claims are enabled thus involves, *inter alia*, the level of skill in the art and the amount of direction provided by the inventor. *In re Wands*, 858 F.2d 731, 737 (Fed. Cir. 1988). Whether the specification provides sufficient disclosure to enable the full scope of a claim is judged from the perspective of one of ordinary skill in the art. M.P.E.P. § 2164.05(b); *see also Falkner v. Inglis*, 448 F.3d 1357, 1365 (Fed. Cir. 2006).

The test for what level of experimentation is undue is "not merely quantitative, since a considerable amount of experimentation is permissible, if it is merely routine, or if the specification in question provides a reasonable amount of guidance with respect to the direction in which the experimentation should proceed." *PPG Industries, Inc. v. Guardian Industries Corp.*, 75 F.3d 1558, 1564 (Fed. Cir. 1996) (quoting with approval *Ex parte Jackson,* 217 USPQ 804, 807 (Bd. Pat. Ap. & Inter. 1982)). Thus, even a large

amount of experimentation is allowed if it would have been routine for one of skill in the art at the time of invention when presented with the guidance of the application. In particular, the presence of actual working examples in a specification supports a finding of enablement. *In re Wands*, 858 F.2d at 740.

The Board has recently issued a precedential decision addressing the amount of disclosure that provides enabling support for claims reciting protein variants. *Ex Parte Kubin*, 83 USPQ.2d 1410 (Bd. Pat. Ap. & Inter. 2007). In *Kubin*, the claim at issue recited polynucleotides encoding polypeptides that were "at least 80% identical" to a given amino acid sequence and that could bind a protein known as CD48. *Id.* at 1415. According to the Board, the specification disclosed two DNA sequences that encoded the reference amino acid sequence and three fusion proteins that encompassed the amino acid sequence, but did not disclose any variants of the nucleotide sequence. *Id.* The Board also found that "the Specification does not disclose which 20% . . . of amino acid residues should be changed to maintain the biological functions" of the polypeptide and that no correlation between the amino acid structure and the ability to bind CD48 was disclosed. *Id.* at 1415-16 (internal quotations omitted).

Despite these findings, the Board ultimately determined that the specification provided enabling support for the full scope of the claim at issue. *Id.* at 1416. In so doing, the Board found that the specification taught the skilled artisan to make the polypeptide variants and to test the ability of the variants to bind CD48. *Id.* The Board further found that methods of making the claimed nucleic acid sequences and methods to screen for their activity were known in the art and disclosed in the application. *Id.* Accordingly, the Board concluded that the amount of experimentation required to make

other sequences within the scope of the claim was within the skill in the art at the time and "thus would have been routine." *Id.* Thus, although the claim at issue in Kubin encompassed a vast number of possible variants, undue experimentation would not have been required for one of skill in the art to practice the invention.

Based on the standard set forth above as held by both the Federal Circuit and the Board and as articulated in the M.P.E.P., Appellant contends that the specification provides sufficient disclosure to allow one of skill in the art to practice the full scope of claims 40-70, 72, 74-77 and 80-91. Arguments are presented below for each group of claims. **Claims argued under different subheadings below do not stand or fall together.**

Claims 40-52, 55-66, 67-70, 72 and 74-77

Claims 40-52, 55-66, 67-70, 72 and 74-77 recite methods to produce glucosamine by fermentation comprising culturing a genetically modified bacterium or yeast to produce glucosamine or glucosamine-6-phosphate and recovering the product. The present claims do not recite specific sequences of glucosamine-6-phosphate synthases, *i.e.*, sequences with specific amino acid mutations. Rather, the claims recite a bacterium or yeast that comprises at least one genetic modification that increases the activity of glucosamine-6-phosphate synthase compared to the unmodified glucosamine-6-phosphate synthase. The genetic modification may be either the expression of a recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase or the modification of the coding region of a gene encoding a bacterial or yeast glucosamine-6-phosphate synthase. Appellant submits that the

specification, combined with the knowledge in the art at the time of invention, provides enabling support for claims 40-52, 55-66, 67-70, 72 and 74-77.

The specification specifically exemplifies the making of bacteria that express a recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase. In particular, Example 2 teaches the production of *E. coli* that overexpress recombinant glucosamine-6-phosphate synthase. Examples 3 and 4 then demonstrate that these bacteria produce an increased amount of glucosamine in fermentation cultures as compared to bacteria that do not express the recombinant synthase. Accordingly, the present specification expressly provides working examples of the claimed methods wherein the genetic modification that increases the activity of glucosamine-6-phosphate synthase is the expression of a recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase.

The inventors also exemplify the production of bacteria that express a recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase that comprises a genetic modification that increases the activity of the synthase compared to the unmodified enzyme. Specifically, Example 5 teaches the use of error-prone PCR to amplify glucosamine-6-phosphate synthase genes while inserting random mutations within the coding sequence of the genes. Example 6 discloses the production of at least three bacterial strains that express glucosamine-6-phosphate synthase genes with a modification of the coding region and that these bacteria produce and increased amount of glucosamine in fermentation cultures as compared to bacteria that do not express the mutant synthase.

The specification further discloses that the invention, while exemplified in bacteria, can be applied to other microorganisms. The amino sugar metabolic pathways and genes and proteins involved therein are known in the art for many bacteria and yeast. Indeed, the Examiner has recognized that bacterial or yeast glucosamine-6-phosphate synthases can be used in the claimed methods. *See, e.g.*, Office Action dated April 21, 2005, at 3-5.

Appellant therefore submits that no undue experimentation would be required to make and use the glucosamine-6-phosphate synthases with the recited genetic modifications, nor to culture bacteria or yeast expressing the same. Indeed, the present inventors have exemplified the production of bacteria expressing these glucosamine-6-phosphate synthases and their use to produce increased amounts of glucosamine. One of skill in the art need only reproduce the detailed instructions specifically set forth in the instant application to practice the full scope of the claimed invention. Moreover, the procedures exemplified in the specification should not present an undue burden for skilled artisans in fields such as molecular biology and microbiology, where the level of skill was recognized to be high at the time of invention. *See. e.g., Kubin* at 1416.

Nonetheless, the Examiner refers to Guo et al. to support his assertion that only a small percentage of possible modifications would increase synthase activity, and the claimed invention thus requires "finding a few mutants within several billion or more." Office Action dated May 2, 2007, at 3. The specification, however, provides a reproducible method to produce mutants and to screen for those that exhibit the claimed properties. The Examiner's reasoning might be on point if the specification did not

provide the means to quickly screen for those modified organisms that fall within the scope of the present claims.

Appellant contends that the Examiner's argument is completely contrary to what has been clearly exemplified and described in the present specification and in subsequent declarations and arguments, and further, goes beyond what is actually required by the present claims. In fact, the inventors have demonstrated that in one routine experiment using the methods of the instant claims, not one, not ten, but **96** microorganisms out of 4368 microorganisms were identified as producers of excess glucosamine compared to the parent strain (see, e.g., Example 5). This is hardly equivalent to needing to find "a few mutants within several billion or more". Thus, the specification clearly demonstrates the production and routine selection of the microorganisms recited in the claims.

Disregarding the express teachings of the specification, the Examiner presents a hypothetical scenario involving a synthase sequence with 25 or 50 mutations. Based on the hypothetical, the Examiner concludes that finding a synthase with 50 mutations would not be possible. Interestingly, the Examiner speculates that "[c]urrent techniques (*i.e.*, high throughput mutagenesis and screening techniques) in the art would allow for finding a few active mutants . . . having 25 mutations" but not 50 mutations. Office Action dated May 2, 2007, at 3.

The instant claims, however, do not recite a mutated synthase having 25 or 50 mutations. Rather, the claims only require that the synthase contain a modification sufficient to increase the synthase activity as compared to the unmodified synthase. The inventors have demonstrated that using techniques known in the art at the time of invention (*i.e.*, high throughput PCR mutagenesis and screening techniques), this result is

readily achievable, requires only routine experimentation, and does not take a long time to accomplish. Appellant thus submits that these teachings provide sufficient guidance to carry out the methods recited in the instant claims without undue experimentation.

Moreover, the specification makes clear that it is not necessary to know where to modify the sequence in order to produce the recited microorganisms and use them in the claimed invention. However, if one wishes to determine the identity of the mutation after the microorganism is identified, this may be accomplished by routine sequencing. The Examiner continues to argue that it is necessary to know what domains and motifs within the amino acid sequence of the *E. coli* glucosamine-6-phosphate synthase can be modified to make a synthase with increased activity. However, the specification, the Declaration of Dr. Deng, and the Declaration of Dr. Demain provide substantial evidence that this is simply not an accurate statement. The Examiner has not provided a specific rebuttal of these arguments or the evidence presented in the Appendix, particularly with respect to the Declaration of Dr. Demain, which includes significant evidence that knowledge of the specific domains and motifs within a sequence are not necessary to routinely make and use genetically modified microorganisms as claimed. The Examiner has not provided clear and convincing proof that undue experimentation would be required to make and use the invention in view of the substantial evidence presented in the record to the contrary.

Applicants note again that the present claims do not recite specific sequences of glucosamine-6-phosphate synthases or sequences with specific amino acid mutations. Instead, the claims recite bacteria or yeast that express modified glucosamine-6-phosphate synthases. As in *Kubin*, the specification teaches the skilled artisan how to

make the modified glucosamine-6-phosphate synthases, how to make bacteria or yeast that express the modified glucosamine-6-phosphate synthases, and how to screen the bacteria or yeast for the increased production of glucosamine. Further, methods for making the modified synthases and methods to screen for the activity of microorganisms expressing the same were either known in the art at the time of invention or are disclosed in the application. No undue experimentation is required to practice the invention.

Combining the teachings of the specification with the knowledge in the art at the priority date of the application, one of skill in the art could readily make and use the full scope of the claimed invention without undue experimentation. Appellant therefore respectfully requests that the Board direct the Examiner to withdraw the rejection of claims 40-52, 55-66, 67-70, 72 and 74-77 under 35 U.S.C. § 112, first paragraph.

Claims 72 and 82-86

In addition to the arguments set forth above for claims 40-52, 55-66, 67-70, 72 and 74-77, Appellant further contends that the specification provides sufficient disclosure to allow one of skill in the art to practice the full scope of claims 72 and 82-86.

Claims 72 and 82-86 recite methods to produce glucosamine similar to those recited in claims 40-52, 55-66, 67-70, 72 and 74-77. Claims 72 and 82-86 further specify that the bacterium or yeast express a recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase with a genetic modification that reduces the glucosamine-6-phosphate product inhibition of the glucosamine-6-phosphate synthase compared to the unmodified glucosamine-6-phosphate synthase.

As discussed in detail above, the specification exemplifies the making of bacteria that express a recombinant nucleic acid molecule encoding a bacterial or yeast

glucosamine-6-phosphate synthase, as recited in claims 72 and 82-86. Example 2 teaches the production of a recombinant DNA molecule encoding a glucosamine-6-phosphate synthase and the overexpression of the recombinant glucosamine-6-phosphate synthase in *E. coli*. Example 2 also demonstrates the integration of the recombinant nucleic acid molecule into the genome of the bacteria. Examples 3 and 4 demonstrate that these bacteria produce an increased amount of glucosamine in fermentation cultures as compared to bacteria that do not express the recombinant synthase.

The inventors also exemplify the production of bacteria that express a recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase that comprises a genetic modification that reduces the glucosamine-6-phosphate product inhibition of the synthase compared to the unmodified enzyme. As discussed above, Example 5 teaches the use of error-prone PCR to amplify glucosamine-6-phosphate synthase genes while inserting random mutations within the coding sequence of the genes. Example 5 also demonstrates the production of six bacterial strains that express mutated glucosamine-6-phosphate synthase genes. Of the six strains, three were then shown to express synthases that were significantly less sensitive to inhibition by the product glucosamine-6-phosphate than control strains using a spectrophotometric enzyme activity assay. Accordingly, the present specification exemplifies, in one routine experiment following the teachings of the application, the production of three separate bacterial strains that meet the limitations of claims 72 and 82-86.

The specification further discloses that the invention, while exemplified in bacteria, can be applied to other microorganisms. The amino sugar metabolic pathways and genes and proteins involved therein are known in the art for many bacteria and yeast.

Indeed, the Examiner has recognized that bacterial or yeast glucosamine-6-phosphate synthases can be used in the claimed methods. *See, e.g.*, Office Action dated April 21, 2005, at 3-5.

Accordingly, no undue experimentation is required for one of skill in the art to practice the full scope of claims 72 and 82-86. Appellant thus requests that the Board direct the Examiner to withdraw the rejection as applied to at least these claims.

Claims 53, 54, 80, 81 and 87-91

In addition to the arguments set forth above for claims 40-52, 55-66, 67-70, 72 and 74-77, Appellant further contends that the specification provides sufficient disclosure to allow one of skill in the art to practice the full scope of claims 53, 54, 80, 81 and 87-91.

Claims 53, 54, 80, 81 and 87-91 recite methods to produce glucosamine similar to those recited in claims 40-52, 55-66, 67-70, 72 and 74-77. Claims 53, 54, 80, 81 and 87-91, however, do not recite the genetic modification language that forms the basis for the enablement rejections. Specifically, the claims do not recite the modification of the coding region of a gene encoding a bacterial or yeast glucosamine-6-phosphate synthase. Rather, claims 53, 54, 80, 81 and 87-91 specify that the genetic modification comprises the transformation of the bacterium or yeast with a recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase that has glucosamine-6-phosphate synthase enzymatic activity.

As discussed in detail above, the specification exemplifies the making of bacteria that express a recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase, as recited in claims 53, 54, 80, 81 and 87-91.

Example 2 teaches the production of a recombinant DNA molecule encoding a glucosamine-6-phosphate synthase and the overexpression of the recombinant glucosamine-6-phosphate synthase in *E. coli*. Example 2 also demonstrates the integration of the recombinant nucleic acid molecule into the genome of the bacteria. Examples 3 and 4 demonstrate that these bacteria produce an increased amount of glucosamine in fermentation cultures as compared to bacteria that do not express the recombinant synthase.

The specification further discloses that the invention, while exemplified in bacteria, can be applied to other microorganisms. The amino sugar metabolic pathways and genes and proteins involved therein are known in the art for many bacteria and yeast. Indeed, the Examiner has recognized that bacterial or yeast glucosamine-6-phosphate synthases can be used in the claimed methods. *See, e.g.*, Office Action dated April 21, 2005, at 3-5.

Accordingly, no undue experimentation is required for one of skill in the art to practice the full scope of claims 53, 54, 80, 81 and 87-91. Appellant thus requests that the Board direct the Examiner to withdraw the rejection as applied to at least these claims.

B. Dutka-Malen et al. Does Not Teach or Suggest Each Element of Claims 40, 53, 55, 62-64, 71, 75, 76-79 and 87-89

Claims 40, 53, 55, 62-64, 71, 75, 76-79 and 87-89 have been finally rejected under 35 U.S.C. § 102(b) as being anticipated by Dutka-Malen et al. ("Dutka"). The rejection reasons that "[s]ince the process steps taught by Dutka-Malen et al. are the same as the recited process steps, then the process would produce glucosamine and the

harvesting of the cells by centrifugation would result in the recovery of the produced glucosamine in the remaining culture media." Office Action dated May 2, 2007, at 4.

A *prima facie* case of anticipation requires that a single publication teach, either expressly or inherently, each and every element or limitation of the claim, including any functional limitations. M.P.E.P. § 2131. Contrary to the reasoning of the rejection, the processes of the instant claims and Dutka are neither the same nor sufficiently similar to support a case of anticipation. For the reasons set forth below, Dutka fails to teach or suggest each element of the claims, and thus cannot possibly anticipate the claims.

Appellant submits that Dutka does not teach that a step of culturing produces and accumulates a product selected from the group consisting of glucosamine-6-phosphate and glucosamine, as recited in the instant claims. At best, Dutka only teaches a molecule encoding the wild-type glucosamine-6-phosphate synthase. Dutka does not measure any glucosamine produced by the microorganism, and does not provide any evidence that the expression of their recombinant nucleic acid molecule increases production of glucosamine by the microorganism. Dutka does not teach or suggest even attempting to *detect* whether the *E. coli* actually produce glucosamine-6-phosphate or glucosamine, let alone the *recovery* of either product.

Moreover, Dutka does not teach a method for producing glucosamine-6-phosphate or glucosamine, let alone the methods recited in the instant claims. The Examiner has repeatedly admitted this fact during the prosecution of the application. Office Action dated May 26, 2000, at 2, 4, and 5. Indeed, Dutka is directed to the cloning of the *glmS* gene and the investigation of the catalytic properties of the enzyme through *in vitro* methodology (See page 288, col. 1, second full paragraph). In short, Dutka is

devoid of any suggestion of a method to produce glucosamine-6-phosphate or glucosamine by fermentation and/or to recover such products from a fermentation medium. Each of these elements is recited in the instant claims.

For at least the reasons discussed above, Dutka does not teach each and every element of the pending claims, and thus cannot expressly anticipate the claims. While the nature of the anticipation rejection is not entirely clear, it is possible that the Examiner is asserting that the process disclosed in Dutka inherently anticipates claims 40, 53, 55, 62-64, 71, 75, 76-79 and 87-89. Appellant contends that the standard for inherency is not met by the disclosure of Dutka and that the Examiner has failed to provide a sufficient basis in fact and/or technical reasoning to support inherent anticipation.

While inherent or implicit disclosures of references may be used to support rejections under 35 U.S.C. § 102, the fact that a certain result or characteristic may occur or be present in the prior art is not sufficient to establish the inherency of that result or characteristic. M.P.E.P. § 2112 (IV) (citing *In re Rijckaert*, 9 F.3d 1531, 1534 (Fed. Cir. 1993)). Rather, the extrinsic evidence must make clear that the characteristic is necessarily and inevitably present in the thing described in the reference, and that it would be so recognized by persons of ordinary skill. *In re Robertson*, 169 F.3d 743, 745 (Fed. Cir. 1999). Inherency, therefore, may not be established by probabilities or possibilities, but only by a characteristic that is always present in a product or occurs each and every time a process is carried out. *Id.* The burden falls on the Examiner to provide a basis in fact and/or technical reasoning to support the assertion that the inherent property necessarily flows from the prior art. *Ex Parte Levy*, 17 USPQ.2d 1461, 1464 (Bd. Pat. App. & Inter. 1990).

Under this standard, Dutka would anticipate the pending claims if, and only if, carrying out the process of Dutka would always result in the production and recovery of glucosamine-6-phosphate or glucosamine from a fermentation culture. However, as discussed above, Dutka does not teach the production of these products nor their recovery from a fermentation culture. Indeed, the only "product" that is produced and partially recovered during the process of Dutka is the synthase itself. The crude synthase extract or semi-purified enzyme is then used in an *in vitro* enzyme activity assay in which additional substrates are added to the reaction.

One of skill in the art knows that the enzyme purification protocol disclosed in Dutka does not always lead to the production of glucosamine-6-phosphate or glucosamine. The skilled artisan is also aware that carrying out the culture process of Dutka would not ever result in the recovery of glucosamine-6-phosphate or glucosamine without the additional *in vitro* assay using semi-purified synthase. In view of this and the knowledge in the prior art, one of skill in the art would clearly recognize that the process of Dutka does not result in the production and recovery of glucosamine-6-phosphate or glucosamine from a fermentation culture. That the process of Dutka could possibly lead to the production and recovery of glucosamine-6-phosphate or glucosamine, after modifications consistent with the teachings of the instant application, does not establish a case of inherency. Likewise, the Examiner's vague assertions that the processes of the instant claims and Dutka "are the same" or that "centrifugation would result in the recovery of the produced glucosamine" lack a basis in fact and are thus insufficient to support inherent anticipation. Office Action dated May 2, 2007, at 4.

Therefore, Appellant respectfully submits that one of skill in the art would know that the procedures disclosed in Dutka would not necessarily and inevitably lead to the production and recovery of glucosamine-6-phosphate or glucosamine from a fermentation culture. Because the pending claims each recite these elements, Dutka cannot anticipate the pending claims either expressly or inherently.

Thus, at least because Dutka does not teach or suggest a method to produce glucosamine by fermentation according to the steps recited in the instant claims, nor the recovery of either glucosamine-6-phosphate or glucosamine from the culture, Dutka does not teach each and every element set forth in the claims. Accordingly, Dutka does not anticipate claims 40, 53, 55, 62-64, 71, 75, 76-79 and 87-89.

In view of the arguments presented above, Appellant respectfully requests that the Board direct the Examiner to withdraw the rejection of claims 40, 53, 55, 62-64, 71, 75, 76-79 and 87-89 under 35 U.S.C. §102(b).

VIII. SIGNATURE OF APPELLANT'S REPRESENTATIVE

Correspondence related to this Appeal Brief should be directed to the undersigned agent, who may also be contacted at (303) 863-9700.

Respectfully submitted,

SHERIDAN ROSS P.C.

Dated: March 28, 2008

By: /John C. Stolpa/
John C. Stolpa
Registration No. 57,632
1560 Broadway, Suite 1200
Denver, CO 80202-5141
(303) 863-9700

IX. CLAIMS APPENDIX

The text of the claims involved in this appeal:

40. A method to produce glucosamine by fermentation, comprising:

a) culturing in a fermentation medium comprising assimilable sources of carbon, nitrogen and phosphate, a bacterium or yeast which comprises at least one genetic modification that increases the activity of glucosamine-6-phosphate synthase compared to the unmodified glucosamine-6-phosphate synthase, wherein said genetic modification is selected from the group consisting of:

i) transformation of said microorganism with a recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase, wherein said synthase has glucosamine-6-phosphate synthase activity; and

ii) genetic modification of the coding region of a gene encoding a bacterial or yeast glucosamine-6-phosphate synthase, wherein said genetic modification results in at least one nucleic acid modification selected from the group consisting of deletion, insertion, and substitution of at least one nucleotide of said coding region of said gene, wherein said at least one nucleotide modification results in increased glucosamine-6-phosphate synthase activity compared to the unmodified glucosamine-6-phosphate synthase;

wherein said step of culturing produces and accumulates a product selected from the group consisting of glucosamine-6-phosphate and glucosamine from said bacterium or yeast; and

b) recovering and purifying said product.

41. The method of Claim 40, wherein said glucosamine-6-phosphate is intracellular and said glucosamine is extracellular, wherein said step of recovering comprises a recovering step selected from the group consisting of recovering said glucosamine-6-phosphate from said bacterium or yeast, recovering said glucosamine from said fermentation medium, and a combination thereof.

42. The method of Claim 40, wherein said product is intracellular glucosamine-6-phosphate and said step of recovering comprises isolating said glucosamine-6-phosphate from said bacterium or yeast.

43. The method of Claim 40, wherein said product is intracellular glucosamine-6-phosphate and said step of recovering further comprises dephosphorylating said glucosamine-6-phosphate to produce glucosamine.

44. The method of Claim 40, wherein said step of culturing comprises maintaining said source of carbon at a concentration of from about 0.5% to about 5% in said fermentation medium.

45. The method of Claim 40, wherein said step of culturing is performed at a temperature from about 30°C to about 40°C.

46. The method of Claim 40, wherein said step of culturing is performed at about 30°C.

47. The method of Claim 40, wherein said step of culturing is performed in a fermentor.

48. The method of Claim 47, wherein said step of culturing is performed under conditions wherein glucose is added to said fermentation medium at a rate in which glucose accumulation in said fermentation medium is undetectable.

49. The method of Claim 47, wherein said step of culturing is performed so that an excess of glucose is maintained.

50. The method of Claim 40, wherein said step of culturing produces and accumulates at least about 21 mg/L of said product.

51. The method of Claim 40, wherein said step of culturing produces and accumulates at least about 1 g/L of said product.

52. The method of Claim 40, wherein said step of culturing produces and accumulates at least about 5 g/L of said product.

53. The method of Claim 40, wherein said genetic modification comprises transformation of said bacterium or yeast with a recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase that has glucosamine-6-phosphate synthase enzymatic activity, wherein said recombinant nucleic acid molecule is operatively linked to a transcription control sequence.

54. The method of Claim 53, wherein said recombinant nucleic acid molecule is integrated into the genome of said bacterium or yeast.

55. The method of Claim 53, wherein said recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase comprises a genetic modification which increases the activity of said glucosamine-6-phosphate synthase compared to the unmodified glucosamine-6-phosphate synthase.

56. The method of Claim 40, wherein said recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase or said gene encoding said a bacterial or yeast glucosamine-6-phosphate synthase comprises a genetic modification which reduces glucosamine-6-phosphate product inhibition of said glucosamine-6-phosphate synthase compared to the unmodified glucosamine-6-phosphate synthase.

57. The method of Claim 40, wherein said bacterium or yeast has at least one additional genetic modification in a gene encoding a protein selected from the group consisting of *N*-acetylglucosamine-6-phosphate deacetylase, glucosamine-6-phosphate deaminase, *N*-acetyl-glucosamine-specific enzyme II^{Nag}, phosphoglucosamine mutase, glucosamine-1-phosphate acetyltransferase-*N*-acetylglucosamine-1-phosphate uridyltransferase, phosphofructokinase, Enzyme II^{Glc} of the PEP:glucose PTS, and EIIM,P/III^{Man} of the PEP:mannose PTS, wherein said genetic modification decreases the activity of said protein compared to the unmodified protein.

58. The method of Claim 40, wherein said bacterium or yeast has at least one additional genetic modification in a gene encoding a phosphatase, wherein said genetic modification increases the activity of said phosphatase compared to the unmodified phosphatase.

59. The method of Claim 40, wherein said bacterium or yeast has additional modifications in genes encoding the following proteins: *N*-acetylglucosamine-6-phosphate deacetylase, glucosamine-6-phosphate deaminase and *N*-acetyl-glucosamine-specific enzyme II^{Nag};

wherein said genetic modification decreases the activity of said proteins compared to the unmodified proteins.

60. The method of Claim 40, wherein said bacterium or yeast has additional modifications in genes encoding *N*-acetylglucosamine-6-phosphate deacetylase and glucosamine-6-phosphate deaminase;

wherein said genetic modification decreases the activity of said proteins compared to the unmodified proteins.

61. The method of Claim 60, wherein said genetic modification is a deletion of at least a portion of said genes.

62. The method of Claim 40, wherein said bacterium or yeast is a bacterium.

63. The method of Claim 62, wherein said bacterium is a bacterium of the genus *Escherichia*.

64. The method of Claim 63, wherein said bacterium is *Escherichia coli*.

65. The method of Claim 64, wherein said bacterium comprises at least one additional genetic modification which is a mutation in an *Escherichia coli* gene selected from the group consisting of *nagA*, *nagB*, *nagC*, *nagD*, *nagE*, *manXYZ*, *glmM*, *pfkB*, *pfkA*, *glmU*, *glmS*, *ptsG* and a phosphatase gene, wherein said genetic modification decreases the activity of a protein encoded by said gene compared to the unmodified protein.

66. The method of Claim 40, wherein said bacterium or yeast is a yeast.

67. A recombinant bacterium or yeast for producing glucosamine by a biosynthetic process, said bacterium or yeast being transformed with a recombinant nucleic acid molecule comprising a nucleic acid sequence encoding a bacterial or yeast glucosamine-6-phosphate synthase, said nucleic acid sequence being operatively linked to a transcription control sequence and comprising, in the coding region of the nucleic acid sequence, a genetic modification which increases the activity of said glucosamine-6-phosphate synthase compared to the unmodified glucosamine-6-phosphate synthase;

wherein expression of said nucleic acid sequence increases production of glucosamine by said bacterium or yeast.

68. The recombinant bacterium or yeast of Claim 67, wherein said bacterium or yeast has at least one additional genetic modification in a gene encoding a protein selected from the group consisting of *N*-acetylglucosamine-6-phosphate deacetylase, glucosamine-6-phosphate deaminase, *N*-acetyl-glucosamine-specific enzyme II^{Nag}, phosphoglucosamine mutase, glucosamine-1-phosphate acetyltransferase-*N*-acetylglucosamine-1-phosphate uridyltransferase, phosphofructokinase, Enzyme II^{Glc} of the PEP:glucose PTS, and EIIM,P/III^{Man} of the PEP:mannose PTS, wherein said genetic modification decreases the activity of said protein compared to the unmodified protein.

69. The recombinant bacterium or yeast of Claim 67, wherein said bacterium or yeast has at least one additional genetic modification in a gene encoding a

phosphatase, wherein said genetic modification increases the activity of said phosphatase compared to the unmodified phosphatase.

70. A method to produce glucosamine by fermentation, comprising:

a) culturing in a fermentation medium comprising assimilable sources of carbon, nitrogen and phosphate, a bacterium or yeast having at least one genetic modification that increases the activity of glucosamine-6-phosphate synthase compared to the unmodified glucosamine-6-phosphate synthase, wherein said genetic modification is selected from the group consisting of:

i) transformation of said bacterium or yeast with a recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase which has glucosamine-6-phosphate synthase activity; and

ii) genetic modification of a gene encoding a bacterial or yeast glucosamine-6-phosphate synthase that increases the activity of said glucosamine-6-phosphate synthase compared to the unmodified glucosamine-6-phosphate synthase, and wherein said genetically modified bacterium or yeast is produced by a process comprising the steps of:

(1) generating modifications in an isolated nucleic acid molecule comprising a nucleic acid sequence encoding a bacterial or yeast glucosamine-6-phosphate synthase to create a plurality of modified nucleic acid sequences;

(2) transforming bacteria or yeast with said modified nucleic acid sequences to produce genetically modified bacteria or yeast;

(3) screening said genetically modified bacteria or yeast for glucosamine-6-phosphate synthase activity; and,

(4) selecting said genetically modified bacteria or yeast which have increased glucosamine-6-phosphate synthase activity compared to the unmodified glucosamine-6-phosphate synthase;

wherein said step of culturing produces a product selected from the group consisting of glucosamine-6-phosphate and glucosamine from said microorganism; and,

b) recovering said product.

71. A method to produce glucosamine by fermentation, comprising:

a) culturing in a fermentation medium comprising assimilable sources of carbon, nitrogen and phosphate, a bacterium or yeast that expresses a recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase, wherein said synthase has glucosamine-6-phosphate synthase enzymatic activity, and wherein said step of culturing produces and accumulates a product selected from the group consisting of glucosamine-6-phosphate and glucosamine from said bacterium or yeast; and

b) recovering and purifying said product.

72. The method of Claim 71, wherein said glucosamine-6-phosphate synthase comprises a genetic modification which reduces glucosamine-6-phosphate product inhibition of said glucosamine-6-phosphate synthase compared to the unmodified glucosamine-6-phosphate synthase.

74. The method of Claim 40, wherein the glucosamine-6-phosphate synthase is a yeast glucosamine-6-phosphate synthase.

75. The method of Claim 40, wherein the glucosamine-6-phosphate synthase is a bacterial glucosamine-6-phosphate synthase.

76. The method of Claim 40, wherein said product is extracellular glucosamine-6-phosphate and said step of recovering comprises recovering said glucosamine-6-phosphate from said fermentation medium.

77. The method of Claim 75, wherein the bacterial glucosamine-6-phosphate synthase is from a bacterium of the genus *Escherichia*.

78. The method of Claim 71, wherein said glucosamine-6-phosphate synthase is a bacterial glucosamine-6-phosphate synthase.

79. The method of Claim 78, wherein said bacterial glucosamine-6-phosphate synthase is from a bacterium of the genus *Escherichia*.

80. The method of Claim 71, wherein the bacterium or yeast further comprises a partial or complete deletion of a gene encoding a protein selected from the group consisting of *N*-acetylglucosamine-6-phosphate deacetylase, glucosamine-6-phosphate deaminase, *N*-acetyl-glucosamine-specific enzyme II^{Nag}, phosphoglucosamine mutase, glucosamine-1-phosphate acetyltransferase-*N*-acetylglucosamine-1-phosphate uridyltransferase, phosphofructokinase, Enzyme II^{Glc} of the PEP:glucose PTS, and EIIM,P/III^{Man} of the PEP:mannose PTS, wherein said partial or complete deletion decreases the activity of said protein compared to the unmodified protein.

81. The method of Claim 80, wherein the bacterium or yeast is a bacterium.

82. A method to produce glucosamine by fermentation, comprising:

a) culturing in a fermentation medium comprising assimilable sources of carbon, nitrogen and phosphate, a bacterium or yeast that expresses a recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase, wherein said synthase has glucosamine-6-phosphate synthase enzymatic activity and comprises a genetic modification that reduces the glucosamine-6-phosphate product inhibition of said glucosamine-6-phosphate synthase compared to the unmodified glucosamine-6-phosphate synthase, and wherein said step of culturing produces and accumulates a product selected from the group consisting of glucosamine-6-phosphate and glucosamine from said bacterium or yeast; and

b) recovering and purifying said product.

83. The method of Claim 82, wherein the bacterium or yeast is a bacterium.

84. The method of Claim 83, wherein the bacterial or yeast glucosamine-6-phosphate synthase is a bacterial glucosamine-6-phosphate synthase.

85. The method of Claim 82, wherein the bacterium or yeast is a yeast.

86. The method of Claim 85, wherein the bacterial or yeast glucosamine-6-phosphate synthase is a yeast glucosamine-6-phosphate synthase.

87. A method to produce glucosamine by fermentation, comprising:

a) culturing in a fermentation medium comprising assimilable sources of carbon, nitrogen and phosphate, a bacterium or yeast that expresses a recombinant nucleic acid molecule encoding a bacterial or yeast glucosamine-6-phosphate synthase, wherein said synthase has glucosamine-6-phosphate synthase enzymatic activity, and wherein said step of culturing produces and accumulates a product selected from the group consisting of glucosamine-6-phosphate and glucosamine from said bacterium or yeast; and

b) recovering and purifying said product.

88. The method of Claim 87, wherein the bacterium or yeast is a bacterium.

89. The method of Claim 88, wherein the bacterial or yeast glucosamine-6-phosphate synthase is a bacterial glucosamine-6-phosphate synthase.

90. The method of Claim 87, wherein the bacterium or yeast is a yeast.

91. The method of Claim 90, wherein the bacterial or yeast glucosamine-6-phosphate synthase is a yeast glucosamine-6-phosphate synthase.

X. RELATED PROCEEDINGS APPENDIX

None.

XI. EVIDENCE APPENDIX

1. The Declaration of Dr. Ming-De Deng and supporting figures are attached hereto as Appendix A. This Declaration was submitted on February 19, 2002, and entered into the record by the Examiner on June 30, 2004 (Office Action dated June 30, 2004, at page 4).

2. The Declaration of Dr. Arnold L. Demain and supporting references are attached hereto as Appendix B. This Declaration was submitted on April 21, 2006, and entered into the record by the Examiner on August 14, 2006 (Office Action dated August 14, 2006, at pages 2 and 4).

APPENDIX A

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In Re the Application of:	Group Art Unit: 1652
BERRY et al.	Examiner: Fronda, C.
Serial No.: 09/341,600	DECLARATION OF MING-DE DENG (Under 37 CFR 1.132)
Filed: September 15, 1999	
Atty. File No.: 3161-18-PUS	
For: "PROCESS FOR PRODUCTION OF N-GLUCOSAMINE"	

CERTIFICATE OF FACSIMILE TRANSMISSION

I hereby certify that this paper is being transmitted via facsimile to the United States Patent and Trademark Office on the date shown below.

CONSTANCE ROBNETT
Print or Type Name

Signature: Constance Robnett

Date: 19 February 2002

Assistant Commissioner for Patents
Washington, D.C. 20231

Dear Sir:

I, Ming-De Deng, declare as follows:

1. I am an investigator employed by the Assignee of the above-identified application, I have worked with the inventors of the application with regard to the subject matter claimed therein, and I am familiar with the application.

2. This Declaration is being submitted in conjunction with a Supplemental Amendment and Response to a final Office Action mailed July 17, 2001.

3. The following paragraphs (4) and (5), Tables 1 and 2, and Figures 1 and 2 are presented in support of the claims of the present invention, which are directed to method to produce glucosamine by fermentation. Specifically, the present discussion and data are submitted in support of the phrase "glucosamine-6-phosphate synthase" as used in the claims, in order to demonstrate that glucosamine-6-phosphate synthases from a variety of sources can be used successfully to practice the claimed method.

Specifically, to demonstrate the breadth of the claimed invention, the present inventors and the Assignee of the present application have functionally expressed different glucosamine-6-phosphate synthase genes in *E.coli* and have demonstrated a glucosamine production level in the

modified microorganisms that is higher than in wild type *E. coli*. The genes used in the following experiment were selected from representative species of diverse groups of organisms, such as gram-positive bacterium (*B. subtilis*), and yeast (*C. albicans* and *S. cerevisiae*). The overexpression of an *E. coli* glucosamine-6-phosphate synthase according to the claimed method has previously been demonstrated in the above-identified application. The glucosamine-6-phosphate synthase sequence homology shared by these organisms and *E. coli* is low (all in the range of about 40%), yet overexpression of each of the glucosamine-6-phosphate genes in an *E. coli* host cell resulted in a significant increase in glucosamine production by the host. Therefore, the data described in detail below provides strong evidence that any glucosamine-6-phosphate synthase gene can be used for glucosamine production in any host using the claimed method of the present invention.

4. *Sequence Homology Among Glucosamine-6-Phosphate Synthase Enzymes from Different Organisms*

By the end of 1997, and therefore at the time of the present invention, genes encoding glucosamine-6-phosphate synthase, also known as glucosamine:fructose-6-phosphate amidotransferase, were identified in many different organisms. The bacterial genes are named *glmS* and the eukaryotic homologues are called GFA, GFAT or GFPT. Organisms in which a glucosamine-6-phosphate synthase gene was identified at the time of the invention include Gram negative bacteria, Gram positive bacteria, yeast, nematode, mouse and human.

Representative glucosamine-6-phosphate synthase amino acid sequences were selected from representative organisms and aligned using the *E. coli* GlmS sequence as a reference. The sequence identities of the organisms are shown in Table 1. Also listed are the number of amino acid residues and deduced molecular weight of the selected enzymes. As shown in Table 1, the homology between *E. coli* GlmS and other sequences is in the range from 41.1% to 72.5%. The GFAT sequences in two yeast strains, *Candida albicans* and *Saccharomyces cerevisiae*, share 72.3% identical residues. The human GFAT sequence shows a homology of 42.1%, 57.5% and 99% to *E. coli*, *S. cerevisiae* and mouse sequences, respectively. The *E. coli glmS* nucleotide sequence shares 45 to 66.4% homology with other sequenced glucosamine-6-phosphate synthase genes listed in Table 1.

Despite differences in sequence, all characterized glucosamine-6-phosphate syntheses share many similar features. For example, their K_m to glutamine, K_m to fructose-6-phosphate and optimal pH are very similar. At the amino acid sequence level, seven residues were shown to be important in catalysis: cysteine-1, asparate-29, histidine-86, histidine-97, asparate-123, cysteine-300 and lysine-603 (*E. coli* GlmS residue numbering). These residues are well conserved in sequences from different organisms. The initiator methionine in glucosamine-6-phosphate synthases is removed enzymatically after translation since the residue is not present in the mature proteins purified from *E. coli, T. thermophilus* and rat. The consensus sequence for the N-terminus is Cysteine-Glycine-Isoleucine. In fact, a cysteine is present at the N-terminal extremity of the mature form of all class-II GATase proteins (class of enzymes to which glucosamine-6-phosphate synthases belong). The cysteine residue has been shown to be important for the catalytic mechanism. The prosite PS00443 describes the specific amino acid residue pattern at the N-terminal of mature class-II GATase proteins.

Therefore, even with low sequence homology, there are conserved structural and biochemical features among the glucosamine-6-phosphate synthases that result in similar function of the enzymes.

5. *Expression of Different Glucosamine Synthase Genes for Glucosamine Production in E. coli*

Glucosamine-6-phosphatesynthase genes from bacteria(*glmS*) and yeast (GFA) were cloned and expressed in *E. coli* to demonstrate their utility in glucosamine production as claimed in the present application. The *glmS* and GFA coding sequences were amplified from *Bacillus subtilis, Saccharomyces cerevisiae* and *Candida albicans* by PCR and placed under the T7 promoter control in the expression vectors pET24d(+) or pET23b(+). The constructs were transformed into the *E. coli* strain 7107-17 (DE3) and maintained as free replicating plasmids. Cell cultures of different strains were induced with IPTG and evaluated for protein expression, glucosamine synthase activity and glucosamine production upon IPTG induction.

The *B. subtilis glmS* gene contains an open reading frame of 1803 bp and encodes a protein of about 65 kDa (599 residues, excluding the initiator methionine which is usually removed in the

cells). The *glmS* gene was amplified by PCR from the strains ATCC 23856 and ATCC 23857. PCR products of expected size were ligated into pET24d(+) (Novagen Inc, Wisconsin). The recombinant plasmids were confirmed by restriction analysis and transformed into 7101-17 (DE3), generating *E. coli* strains 7107-24 (*glmS* gene from *B. subtilis* ATCC23856) and 7107-25 (*glmS* gene from *B. subtilis* ATCC23857). As a control, the empty vector pET24d was also transformed into 7101-17(DE3), generating the strain 7107-22.

The *S. cerevisiae* GFA1 open reading frame has 2154 bp and encodes a peptide of 716 residues (excluding the initiator methionine). The protein size predicted from the sequence is about 80 kDa. There are no introns in the GFA1 gene sequence. Therefore, the gene was amplified from genomic DNA prepared from the strain *S. cerevisiae* S288C. The PCR product of about 2.2 kb was cloned into pCR-Script Amp SK(+). Recombinant plasmids were confirmed by restriction enzyme digestions. The *S. cerevisiae* GFA1 fragment was isolated by digestion with *Eco*R I and *Bsa* I and ligated into the *Eco*R I and *Nco* I sites of pET24d(+). The recombinant plasmid was confirmed by restriction analysis and transformed into 7101-17(DE3), generating the *E. coli* strain 7107-101.

The *C. albicans* GFA1 gene is free of introns and its 2142-bp open reading frame encodes a peptide of about 80 kDa (712 residues, excluding the initiator methionine). The GFA1 coding sequence was amplified from the strain ATCC10261 by PCR. The PCR product was cloned into the vector pMOSBlue (Amersham Pharmacia Biotech, New Jersey) and recombinant plasmids were confirmed by restriction enzyme digestion. The *Bsa* I-*Xho* I fragment was isolated and ligated into pET24d(+) prepared by digestion with *Nco* I and *Xho* I. The resulting plasmid was transformed into the host 7101-17 (DE3), generating the *E. coli* strain 7107-23.

The *C. albicans* GFA1 gene was also cloned into the expression vector pET23b (Novagen Inc). Unlike pET24d, this vector does not contain a *lacI* repressor gene and it does not have a *lac* operator sequence downstream from the T7 promoter. The use of this vector often results in a higher recombinant protein expression. The recombinant plasmid was confirmed by restriction analysis and transformed into the expression host 7101-17 (DE3), generating *E. coli* strains 7107-58 and 7107-59. As a control, the empty vector pET23b was also transformed into 7101-17(DE3), generating the strain 7107-57.

Strains transformed with pET vectors containing different *glmS* and GFA 1 genes described above were grown in LB medium and induced with 1 mM IPTG to demonstrate GlmS and GFA1 protein expression. As a negative control, cells with the empty pET24d vector were also grown and analyzed. For comparison, *E. coli* cells with the wild-type *E. coli glmS* gene and mutant *glmS*54* gene driven by the T7 promoter and integrated in the chromosome at the *lacZ* site were also grown and analyzed.

To investigate the expression level of the enzymes, SDS-PAGE was carried out by following standard methods. When the *T7-E. coli glmS* expression cassette was carried in pET plasmids or integrated in the chromosome, the GlmS protein was expressed at very high levels (Fig. 1, lane 2 and 3). Cells hosting the plasmids pET24d/*T7-B. subtilis glmS* over-expressed a protein of about 65 kDa, the expected size of the GlmS protein (lane 4 and 5). The expression level was comparable to the cells expressing the *E. coli glmS* gene contained in pET plasmids. Cells hosting the *S. cerevisiae* GFA gene showed an overexpressed protein band of the expected size for the yeast GFA1 protein (80 kDa, lane 6). In the strain 7107-23 containing the T7-*C. albicans* GFA1 expression cassette (lane 7), the synthesis of the 80-kDa protein band was not apparent when compared to the strain with the empty vector (lane 8). However, the GFA1 band was overexpressed in the strains 7107-58 and 7107-59 containing the *C. albicans* GFA1 gene carried by the vector pET23b (Fig. 2). The expression level was higher than in the strain 7107-23 with the pET24d-based vector. The use of alternative codons for Leu 29 and Ala 655 did not affect *C. albicans* GFA1 protein expression in *E. coli*. Expression levels of the yeast GFA genes in *E. coli* were low as compared to bacterial *glmS* genes. This is commonly observed when attempting to express eukaryotic genes in *E. coli* hosts. However, as discussed below, cultures expressing the yeast GFA genes still produced significantly increased amounts of glucosamine as compared to the controls.

For measurement of enzyme activity and glucosamine production, different strains were grown in M9A medium. Data from representative experiments are shown in Table 2. The enzyme activity was readily detectable in *E. coli* cells expressing the *B. subtilis glmS* genes. The activity level was comparable to the cells with a construct containing the *E. coli glmS* and *E. coli glmS*54* mutant

gene. A trace amount of enzyme activity could be detected in cells hosting the yeast GFA 1 genes, and this was correlated with the lower protein expression levels as discussed above.

Only a very low level of glucosamine was produced and secreted into the culture medium of 7101-17 (DE3) cells transformed with an empty vector pET24d (Table 1). Expression of a bacterial *glmS* gene (*E. coli glmS* or *B. subtilis glmS*) resulted in greater than 50-fold increase in glucosamine production. A several-fold increase in glucosamine level was also observed in the cultures expressing yeast GFA1 genes, demonstrating that even at a lower level of enzyme overexpression and activity as compared to the bacterial genes, *significant glucosamine production is achieved.* As compared to pET24d, the use of pET23b led to a higher level of *C. albicans* GFA1 protein and a higher level of glucosamine production. As observed in enzyme activity assays, integration of the *T7-E. coli glmS* expression cassette in the chromosome appeared to be beneficial, as a higher glucosamine level was produced in the strain 2123-12 than in 7107-214. The *E. coli* strain with *E. coli glmS*54* integrated in the chromosome was the highest for glucosamine production when compared to other tested strains.

In summary, these experiments demonstrate that, using the guidance provided in the present application, the method claimed in the above-identified application can be predictably and successfully practiced using glucosamine-6-phosphate synthase from a variety of organism sources.

6. I hereby declare that all statements made herein of my own are true and that all statements made on information and belief are believed to be true; and further that the statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the subject application or any patent issuing therefrom.

Date: Feb 15, 2002 By: M Deng

Ming-De Deng

Table 1. Amino Acid Sequence Comparison of Glucosamine Synthases from Different Organisms*

*: Sequences were published before January, 1998

**: Number of amino acid residues does not include the initiator methionine, which is removed enzymatically after translation.

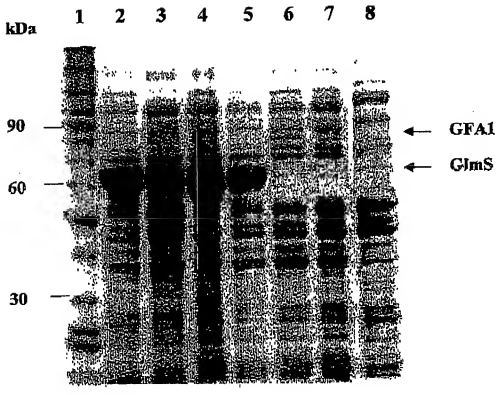
***: Percentage of identical residue was determined by the Method of J Hein, Lasergene software

	Sequence	Number of	Molecular	Identical Residues (%)***											
	Name	Residues**	Weight (kDa)	1	2	3	4	5	6	7	8	9	10	11	12
1	*E. coli* GlmS	608	66.8		72.5	47.7	49.2	44.0	41.1	46.6	42.1	42.1	41.1	41.7	42.1
2	*H. influenzae* GlmS	609	66.7			46.8	49.1	42.8	40.0	45.4	42.0	41.3	40.7	43.0	43.3
3	*R. leguminosarum* GlmS	607	66.7				48.3	44.3	41.2	44.2	38.2	38.7	37.7	39.7	39.5
4	*T. thermophilus* GlmS	603	66.4					50.4	43.9	47.1	39.5	38.6	39.4	40.0	39.8
5	*M. leprae* GlmS	624	67.4						40.1	41.6	37.9	37.8	38.2	38.1	37.8
6	*B. subtilis* GlmS	599	65.2							40.3	36.1	35.6	34.1	35.6	35.9
7	*Synechocystis sp.* GlmS	630	69.5								38.9	37.7	37.5	38.6	38.6
8	*C. albicans* GFA1	712	79.1									72.3	52.7	56.4	56.6
9	*S. cerevisiae* GFA1	716	79.9										53.6	57.3	57.5
10	*C. elegans* GFAT	709	79.2											61.3	61.0
11	Mouse GFAT1	680	76.6												99.0
12	Human GFAT1	680	76.6												

Table 2. Glucosamine synthase activity and glucosamine production in *E. coli* strains expressing different *glmS* and GFA homologues

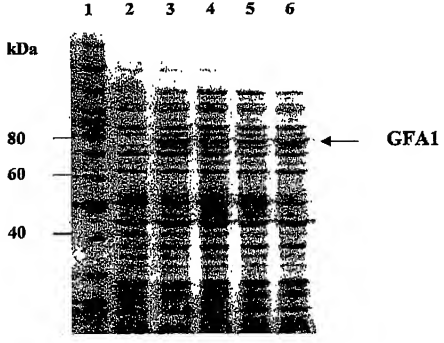
Strain Number	Strain description	Enzyme activity ($nmol.min^{-1}.mg^{-1}$)	Glucosamine ($mg.l^{-1}$)
7107-22	pET24d	trace	5
7107-24	pET24d /T7-*B. subtilis glmS* 23856	637	128
7107-101	pET24d/T7-*S. cerevisea* GFA1	trace	47
7107-23	pET24d/T7-*C. albicans* GFA1	trace	23
7107-58	pET23b/T7-*C. albicans* GFA1	trace	54
7107-60	pET23b/T7-*C. albicans* GFA1-M	trace	58
7107-214	pET24d/T7-*E. coli glmS*	297	37
2123-12	*lacZ::T7-E. coli glmS*	613	75
2123-54	*lacZ::T7-E. coli glmS*54*	803	2,029

Notes: 1) Host cell: *E. coli* 7101-17 (DE3). Genotype: *nag, manXYZ* DE3.
2) Cell culture: 30°C for 26 hrs in shake flasks containing M9A medium supplemented with 7.5 g $(NH_4)_2SO_4$ per liter and 40 g glucose per liter.
3) *C. albicans* GFA1 (M): Leu 29 and Ala 655 codons changed from TTA and GCT to CTA and GCC, respectively.



1: Size Standard
2: 7107-214 (pET24d/T7-*E. coli glmS*)
3: 2123-12 (*lacZ::T7-E. coli glmS*)
4: 7107-24 (pET24d/T7-*B. subtilis glmS* 23856)
5: 7107-25 (pET24d/T7-*B. subtilis glmS* 23857)
6: 7107-101 (pET24d/T7-*S. cerevisiae* GFA1)
7: 7107-23 (pET24d/T7-*C. albicans* GFA1)
8: 7107-22 (pET24d, control)

FIG. 1



1: protein size markers
2: 7107-57(pET23b)
3: 7107-58 (pET23b/T7-*C. albicans* GFA1)
4: 7107-59 (pET23b/T7-*C. albicans* GFA1)
5: 7107-60 (pET23b/T7-*C. albicans* GFA1-M)
6: 7107-61 (pET23b/T7-*C. albicans* GFA1-M)

FIG. 2

APPENDIX B

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In Re the Application of:	)	Group Art Unit: 1652
BERRY et al.	)	Examiner: Fronda, C.
Serial No.: 09/341,600	)	
Filed: September 15, 1999	)	DECLARATION OF DR. ARNOLD L. DEMAIN (Under 37 CFR 1.132)
Atty. File No.: 3161-18-PUS	)	
For: "PROCESS FOR PRODUCTION OF N-GLUCOSAMINE"	)	

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313

Dear Sir:

I, Arnold L. Demain, declare as follows:

1. I am an expert in the field of fermentation biology and strain improvement, and I have worked in this field for more than 50 years. Since 2001, I have been a Research Fellow in Microbial Biochemistry at The Charles A. Dana Research Institute for Scientists Emeriti (R.I.S.E.) at Drew University in Madison, New Jersey. I am also Visiting Professor at Rutgers University. I was formerly Professor of Industrial Microbiology in the Department of Biology at the Massachusetts Institute of Technology (MIT) for 32 years. Prior to MIT, I was employed by Merck & Co., Inc., and was founder and head of the Department of Fermentation Microbiology. I received my Bachelor of Science and Master of Science degrees in bacteriology from Michigan State University in 1949 and 1950, respectively, and my doctorate in microbiology in 1954 from the University of California at Davis and Berkeley. My honors, awards, and honorary degrees are listed in my curriculum vitae, which is attached to this Declaration. I have published over 490 publications, co-authored or co-edited 11 books, and hold 21 U.S. patents. I am also a member of the National Academy of Sciences in the United States, Mexico, and Hungary, and a member of the Board of Governors of the American Academy of Microbiology. This summary

of my experience, in conjunction with the attached curriculum vitae, provides evidence of my extensive experience in the fields of fermentation biology and strain improvement.

2. I have been engaged as an independent third party expert by Arkion Life Science LLC d/b/a/ Bio-Technical Resources Division, of Wilmington, Delaware, which is the Assignee of the above-identified application, to review the file of the above-identified application, and to provide my comments on certain issues in the prosecution of the application. I was compensated for my time on an hourly basis.

3. This Declaration is being submitted in conjunction with an Amendment and Response to an Office Action having a mailing date of April 21, 2005 in the above-identified application.

4. I have reviewed, and I am familiar with, the above-identified application and the presently pending claims, and I have also reviewed several documents from the file history of this application, including the prior Declaration under 37 CFR 1.132 that was submitted by Dr. Ming-De Deng, and the Office Action mailed April 21, 2005. The following discussion provides my comments on the Examiner's rejection of Claims 40-76 under 35 U.S.C. § 112, first paragraph, on the basis of enablement. In particular, the following discussion addresses the Examiner's argument that the specification does not enable the invention that is claimed.

5. Comments on the Examiner's Rejection

Introduction.

The lifeblood of the fermentation industry is the ability to improve microbial strains genetically. This has been practiced since the nineteen forties at a time when penicillin was a wartime necessity. At first, it was done by random mutation, *i.e.*,"brute-force" screening technology in which the wild-type strain, and successive improved strains, were mutagenized by toxic chemicals (such as nitrosoguanidine) or physical agents (such as Xrays or UV). Most of the population was killed and the remaining cells were screened by fermentation in flasks containing a liquid medium for a number of days, and testing the resultant broths for the amount of product made. Such screening could involve as many as 10,000 to 100,000 surviving isolates before an improved strain was isolated.

This tedious procedure was improved over the years by more convenient screening methods in which product formation was first assayed on agar plates, *e.g.*, measuring the

diameter of a clear zone ("halo") around an antibiotic-producing colony or around an agar plug containing antibiotic-producing cells on its surface (the "agar-plug" method), or the color of the colony in the case of pigmented products, or the diameter of a colored halo around a colony producing a compound capable of a chemical reaction producing a color in the presence of a reagent, or the diameter of a cloudy halo around a colony producing a growth-promoting compound.

Another major improvement was the employment of **selection**, *i.e.*, the use of antimetabolites to kill a large part of the population surviving mutagenesis and thus decrease the number of isolates to be assayed by fermentation. In such cases, the antimetabolites were chosen on a rational basis, usually structurally-related analogs of the desired compound, to select those mutants which were most likely improved in production ability.

The molecular biology revolution occurring in the mid-1950's, and the biotechnology revolution of the 1970's allowed a more rational basis for creating genetic diversity than random mutation. This included the transfer of genes between species using plasmid transformation, transposon mutagenesis, protoplast fusion, homologous recombination, polymerase chain reaction (PCR) cloning, gene sequencing, DNA shuffling, whole genome shuffling and other techniques. However, such modern gene manipulation employed for strain improvement never eliminated the need for screening/selection. It merely cut down on the number of isolates needed to be screened from hundreds of thousands or tens of thousands to a few thousand or even less. Thus, screening/selection is an obligatory part of strain improvement, *i.e.*, it is a routine procedure used in various forms by all microbiologists, microbial geneticists, and molecular biologists. Recent improvements in screening technologies have involved miniaturization of flask fermentation into test tubes, or into microtiter plates containing 96 or more wells, followed by assay by high throughput screening techniques. Clearly, screening is still essential and routine to any molecular genetic improvement process.

Comments on Enablement Rejection and Claimed Invention.

I do not agree with the Examiner when he states that routine experimentation in the art does not include making vast numbers of mutants, and screening and selecting said mutants. I also do not agree when the Examiner states that one skilled in the art would require additional guidelines, such as the specific type of genetic modification to perform on the specific proteins

or enzymes of claimed microorganisms, or the amino acid residues which can be modified that lead to the claimed effect, or the gene encoding the enzyme and its biological source. The Examiner is mistaken when he states that without such guidance, the experimentation left to those skilled in the art is undue or is unpredictable. In my mind and experience, it is not undue or unpredictable. Indeed, it is an expected type of effort; it is simple, and will succeed as long as the method of screening/selection is clearly presented, as it has been in the above-identified patent specification. The Examiner is mistaken when he states that the work of the experimenter would include selecting from a variety of genetic modifications such as replacing the wild-type promoter of the glucosamine-6-phosphate synthase, modifying the nucleic acid sequence encoding the synthase, and selecting proteins other than the synthase to modify. Indeed, it is my view that these efforts would be unnecessary to carry out the invention described in the above-identified application.

I agree with the Applicants that the original patent specification describes the development of a strain by both classical strain development and molecular genetic techniques and gives a detailed description and working examples of how to use molecular genetic techniques and classical strain development. I also agree with the Applicants' contention that one skilled in the art could produce strains with increased glucosamine-6-phosphate synthase activity without the knowledge of where in the nucleotide sequence or enzyme sequence the mutation occurs. I further agree with the Applicants when they state that routine experimentation does include making a vast amount of mutants and screening and selecting those which have the desired phenotype of increased glucosamine-6-phosphate synthase activity. I am in agreement with the Applicants' contention that biologists have been producing and selecting desired mutants for many years, and that the predictability of producing an important strain using the guidance in the patent specification is quite high, and that routine experimentation is all that is required to produce, screen and select multiple strains having the recited genetic modification(s). I also agree with the Applicants when they state that the advent of recombinant technology cannot be used by the Examiner to dismiss classical techniques that have been used for years. I also agree with the Applicants' statement that once a beneficial mutation is discovered by genetic techniques including screening/selection, the sequence information can be determined and used,

if desired, to produce the same mutation using molecular biology, but that this prior knowledge is not necessary for one of skill in the art to make and use the present invention.

In my opinion, the screening described in Example 5 of 4368 transformants for glucosamine production is routine in the art and was accomplished in a single experiment. It resulted in the finding of 96 improved producers without any knowledge of the sequence of the enzyme or of the gene involved. Thirty of these 96 strains had the largest halos, and were rated as superior. Six strains were chosen from the superior 30 as the best producers of glucosamine. This simple procedure is predictable with regard to yielding strains meeting the requirements of improved glucosamine production. Of the 6 strains, three had the expected phenotype of reduced feedback inhibition and produced very high levels of glucosamine, as described in Example 6. The number 3 of improved mutant strains was high and such a result was predictable.

I further disagree with the Examiner when he states that the specification forces one skilled in the art to do trial and error experimentation to arrive at a glucosamine-6-phosphate synthase that has increased activity or decreased product inhibition. This is not "trial and error experimentation". It is experimentation following a carefully detailed procedure which yields a desired result.

To support my position, I cite and briefly discuss a number of recent scientific publications in which molecular biology techniques are accompanied by necessary screening/selection techniques.

Literature Cited

Q. Wang et al., Metabolic Engineering of *Torulopsis glabrata* for Improved Pyruvate Production. Enzyme and Microbial Technology 36:832-839, 2005. *T. glabrata* strains were metabolically engineered for increased production of pyruvate. First, *T. glabrata ura* strains (requiring uracil for growth) that were suitable for genetic transformation were isolated by ethyl methansulfonate mutagenesis, and selection by growth in media containing the antimetabolite 5-fluoroorotic acid. Then, the gene encoding PDC was specifically disrupted via homologous recombination with the *Saccharomyces cerevisiae URA3* gene as the selective marker. The disruptants displayed higher pyruvate production (20 grams/liter) and less of the undesirable ethanol production (4.6 grams/liter) than the parental strain (7-8 grams/liter pyruvate and 7.4 grams/liter ethanol). The disruptants

were able to produce in a 52 hour jar fermentation 82 grams/liter. It is clear that in this impressive work, selection was crucial for the process of strain improvement.

K. Stutzman-Engall et al., Semi-synthetic DNA Shuffling of *avaC* Leads to Improved Industrial Scale Production of Doramectin by *Streptomyces avermitilis*. Metabolic Engineering 7:27-37, 2005. The commercial Pfizer product Doramectin™ is made from the avermectin analog doramectin CHC-B1. This analog is made by fermentation with *Streptomyces avermitilis* along with the undesirable analog CHC-B2. To improve the ratio between the preferred CHC-B1 and the undesirable CHC-B2, the biosynthetic gene *avaC* was subjected to iterative rounds of semisynthetic DNA shuffling, a tool of molecular genetics. The resultant library of 5000 shuffled *aveC* variant strains was screened by fermentation in small plates containing 96 wells and assayed by high throughput mass spectrometry to determine ratios. Eighty nine strains were found to have improved ratios. The 89 were then fermented in flasks yielding 8 strains with the highest ratios. The procedure was repeated and the best strain was found to exhibit a 23-fold ratio improvement over the starting strain. This very important result was thus the product of the modern molecular genetic technique of gene shuffling combined with the necessary process of screening.

J. Mampel et al., Single-gene Knockout of a Novel Regulatory Element Confers Ethionine Resistance and Elevates Methionine Production in *Corynebacterium glutamicum*. Applied Microbiology and Biotechnology. 68:228-236, 2005. *C. glutamicum* was mutagenized by the recombinant DNA technique known as transposon mutagenesis. 7000 mutants were screened for rapid growth on agar plates containing the methionine antimetabolite known as DL-ethionine. Transfer of the ethionine resistance trait to another strain inactivated the *NCg12640* gene resulting in 2-fold overproduction of methionine as compared to the original strain. Thus, success was obtained using the genetic technique of transposon mutagenesis coupled to selection with ethionine and screening in flask fermentation.

G. Stephanopoulos et al., Exploiting Biological Complexity for Strain Improvement through Systems Biology. Nature Biotechnology 22:1261-1267, 2004. This review discusses the new concept of systems biology to improve strain improvement. The technology depends

on introducing a variety of perturbations and measuring the system response. They state that "a diverse set of tools has emerged to create gene deletions and amplifications which are used in conjunction with altered environments. Molecular biology advances have made it possible to perform these modifications at will. In addition to gene-specific techniques, several combinatorial tools have been developed, which, when combined with high-throughput screening, allow for randomized gene-expression levels...Advances in high-throughput assays greatly enhance our ability to characterize the cellular phenotype...Applying these tools to a microbial system provides a detailed snapshot of cellular function."

Z. Xu et al., A High-Throughput Method for Screening of Rapamycin-Producing Strains of *Streptomyces hygroscopicus* by Cultivation in 96-Well Microtiter Plates. Biotechnology Letters 27:1135-1140, 2005. The authors state that traditional mutation is still used to improve industrial strains and that new techniques such as genome shuffling still require positive mutants to be screened out from a large number of isolates obtained by recursive protoplast fusion. They further state that a variety of screening methods have been developed for different types of microorganisms including the "agar-plug" method and subsequent modified versions. In this paper, they describe a new screening method involving the culture of mutants on the surface of 96 agar-solidified wells in microtiter plates. By use of this screening technique, they were able to double the production titer of rapamycin.

W.J. Coleman et al., Solid-Phase High-Throughput Screening of Enzyme Variants: Detecting Enhanced Nitrilase Activity. Industrial Biotechnology 1:102-105, 2005. The authors discuss the modern technique of directed evolution of enzymes and state that it consists of the following steps: (1) mutating the gene encoding the enzyme to create a large population of variants; (2) expressing the gene products in a host organism; (3) screening for the desired properties; and (4) retrieving the desired variants. These steps can be repeated until the ideal enzyme variant is obtained. Their technique called Kcat technology monitors enzymatic activity of tens of thousands of enzyme variants simultaneously by acquiring full spectral and/or kinetic information from microcolonies that are simultaneously undergoing a color-forming reaction catalyzed by the enzyme that

the microbial colonies express. Each assay disk contains about 9000 individual microcolonies. Colonies with the highest activity are picked. Using such a screening method, over one million variants can be screened per instrument per day.

To summarize, let me say that that (1) one skilled in the art would not have to know the specific genetic modification(s) to perform on a specific gene or promoter or protein or enzyme to achieve an improved process; (2) routine experimentation in the art includes screening and/or selection; (3) such screening/selection is a necessary part of all strain improvement efforts, even when using the most modern genetic methods to cause mutations in the microbe; (4) such experimentation is not undue or unpredictable, but is simple, rapid, and will succeed provided that the method of screening/selection is clearly stated, as it has been by the Applicants; and (5) the Applicants have devised an elegant procedure to obtain industrially relevant concentrations of a medically useful molecule and deserve patent protection for the process.

6. I hereby declare that all statements made herein of my own are true and that all statements made on information and belief are believed to be true; and further that the statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both under Section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the subject application or any patent issuing therefrom.

March 16, 2006
Date


Arnold L. Demain

DREW UNIVERSITY

ARNOLD L. DEMAIN, Ph. D.
Fellow in Microbial Biochemistry
Charles A. Dana Research Institute (R.I.S.E.)
HS-330
Madison, N.J. 07940, U.S.A.

Tel. 973-408-3937
Fax. 973-408-3504
Email: ademain@drew.edu

A. PROFESSIONAL AND ACADEMIC ESSENTIALS

Education:

1949, 1950	B.S. & MS, Michigan State University Bacteriology
1954	Ph.D., University of California (Davis and Berkeley), Microbiology

Experience:

1954-1964	Merck & Co, Research Microbiologist
1965-1969	Merck & Co, Founder and Head, Department of Fermentation Microbiology
1969-2001	M.I.T., Professor of Industrial Microbiology
2001-present	Drew University, Fellow, Charles A. Dana Institute for Scientists Emeriti

Honors:

1978	Charles Thom Award, Society for Industrial Microbiology
1979	Rubbo Oration Award, Australian Society for Microbiology
1980	First David Perlman Memorial Lecture Award, American Chemical Society
1981	Fellow, American Academy of Microbiology
1985	Honorary Member of the Kitasato Institute, Tokyo
1985	Fellow, Society for Industrial Microbiology
1986	Chair, International Committee on Economic and Applied Microbiology
1987	Senior U.S. Scientist Humboldt Foundation Prize from the Federal Republic of Germany
1987	Fellow of the International Institute of Biotechnology, UK
1988	Microbial Chemistry Medal from Kitasato Institute, Tokyo
1989	Award of the Italian Industrial Pharmaceutical Association
1990	First Hans Knoll Memorial Medal, Academy of Science of GDR
1990	Biotechnology Award, American Society for Microbiology
1990	President, Society for Industrial Microbiology
1993	Honorary Member, Societe Francaise de Microbiologie
1994	American Society for Microbiology Distinguished Service Award
1994	Elected to membership in National Academy of Science, USA
1994	Waksman Outstanding Teaching Award, Society for Industrial Microbiology
1995	Honorary Membership, Society for Actinomycetes Japan (SAJ)
1996	Corresponding Member, Mexican Academy of Sciences
1997	Doctor Honoris Causa, University of Leon, Spain
1997	Marvin J. Johnson Award in Microbial and Biochemical Technology, American Chemical Society
1997	Dedication Award, 5th International Conference on the Biotechnology of Microbial Products: Novel Pharmolological and Agrobiological Activities (Williamsburg, VA)
1997	Special Award of the Society for Fermentation & Bioenginering of Japan
1998	Alice Evans Award, American Society for Microbiology

1998:	Andrew Jackson Moyer Lectureship Award, NCAUR, US Dept. Agriculture, Peoria
1998	G. J. Mendel Honorary Medal of the Czech Academy of Sciences for Merit in the Biological Sciences
1998	Honorary Membership, Czecholovak Society for Microbiology
1999	Doctor Honoris Causa, Ghent University, Belgium
1999	Honorary Membership Award, Northeast Branch of American Society for Microbiology
1999	Medal of the Order of the Rising Sun, Imperial Decoration Award of Japan
2000	Elected to Board of Governors of American Academy of Microbiology
2000	Honorary Doctorate from Technion (Israel Inst. Technology)
2000	Honorary Doctorate from Michigan State University
2000:	Appointed, Advisory Committee of the Korea Research Institute of Bioscience and Biotechnology
2002	Appointed, Advisory Committee of the University of Iowa Center for Biocatalysis and Bioprocessing
2002	Elected to Hungarian Academy of Science
2003	Honorary Doctorate, University of Muenster, Germany
2003	Re-elected to Board of Governors of American Academy of Microbiology
2005	Arima Award in Applied Microbiology from the International Union of Microbiological Sciences (IUMS)
2005	Shanghai Institute of the Pharmaceutical Industry International Achievement Award
2005	SynerZ Lifetime Achievement Award

Editorial Activities:

Editor, Editorial Board, or Advisory Board of 34 journals & book series; 23 in past; 11 currently.

Books:

Co-Editor or co-author of 11 books.

Patents:

21 US Patents.

B. 495 Publications including: Paiva, Demain and Roberts, The immediate precursor of the nitrogen-containing ring of rapamycin is free pipecolic acid. Enzyme Microb. Technol. 15, 581 (1993). Paiva, Roberts and Demain, The cyclohexane moiety of rapamycin is derived from shikimic acid in *Streptomyces hygroscopicus* . J. Ind. Microbiol. 12, 423 (1993). Kobayashi, Romaniec, Barker, Gerngross & Demain, Nucleotide Sequence of Gene *celM* Encoding a New Endoglucanase (CelM) of*Clostridium thermocellum* and Purification of the Enzyme. J. Ferm. Bioeng. 76, 251 (1993). Gerngross, Romaniec, Kobayashi, Huskisson & Demain, Sequencing of a *Clostridium thermocellum* Gene (*cipA*) Encoding the Cellulosomal SL-Protein Reveals an Unusual Degree of Internal Homology. Molec. Microbiol. 8, 325 (1993). Paiva, Demain and Roberts, The immediate precursor of the nitrogen-containing ring of rapamycin is free pipecolic acid. Enzyme Microb. Technol. 15, 581 (1993). Paiva, Roberts and Demain, The cyclohexane moiety of rapamycin is derived from shikimic acid in *Streptomyces hygroscopicus* . J. Ind. Microbiol. 12, 423 (1993). Romaniec, Huskisson, Barker & Demain, Purification and Properties of the *Clostridium thermocellum bglB* Gene Product Expressed in *Escherichia coli*. Enzyme Microb. Technol. 15, 393 (1993). Nochur, Roberts & Demain, True Cellulase Production by *Clostridium thermocellum* Grown on Different Carbon Sources, Biotechnol. Lett. 15, 641 (1993). Lin and Demain, Resting cell studies on formation of water-soluble red pigments by *Monascus* sp. J. Ind. Microbiol. 12, 361 (1993). Demain, Turning Garbage into Motor Fuel: Fanciful Dream or Feasible Scheme? In: Genetics, Biochemistry and Ecology of Lignocellulose Degradation (Shimada et al., eds) pp 573-583, Uni, Tokyo (1994). Lin and Demain, Leucine interference in the production of water-soluble red *Monascus* pigments. Arch. Microbiol. 162, 114 (1994). Lin and Demain, Negative effect of ammonium nitrate as nitrogen source on the production of water-soluble

red pigments by *Monascus* sp. Appl. Microbiol. Biotechnol. 43, 701 (1995). Kruus, Lua, Demain & Wu, The Anchorage Function of CipA (CelL), a Scaffolding Protein of the*Clostridium thermocellum* Cellulosome. Proc. Natl. Acad. Sci., USA 92, 9254 (1995). Cheng, Hauck and Demain, Phosphate, ammonium, magnesium and iron nutrition of *Streptomyces hygroscopicus* with respect to rapamycin production. J. Ind. Microbiol. 14, 424 (1995). Kojima, Cheng, Mohan and Demain, Carbon source nutrition of rapamycin biosynthesis in *Streptomyces hygroscopicus*.. J. Ind. Microbiol. 14, 436 (1995). Cheng, Fang and Demain, Effect of amino acids on rapamycin biosynthesis by *Streptomyces hygroscopicus*. Appl. Microbiol. Biotechnol.43, 1096 (1995). Demain, Microbial secondary metabolism: the new frontier. In: Sekundärmetabolismus bei Mikroorganismen. Beiträge zur Forschung (W. Kuhn and H.-P. Fiedler, eds.) pp. 9-35. Attempto Verlag, Tübingen (1995). Fang and Demain, Dependence of nitrogen- and phosphorus-regulation of β-lactam antibiotic production by Streptomyces clavuligerus on aeration level. J. Indust. Microbiol. 15, 407 (1995). Demain and Fang, Emerging concepts of secondary metabolism in actinomycetes. Actinomycetologica 9, 98 (1995). Fang and Demain, Exogenous shikimic acid stimulates rapamycin biosynthesis in *Streptomyces hygroscopicus*. Folia Microbiol. 40, 607 (1995). Klapatch, Demain and Lynd, Restriction endonuclease activity in *Clostridium thermocellum* and *Clostridium thermosaccharolyticum*. Appl. Microbiol. Biotechnol. 45, 127 (1996). Cochet and Demain, Effect of water activity on production of β-lactam antibiotics by Strepromyces clavuligerus in submerged culture. J. Appl. Bacteriol. 60, 333 (1996). Xiao, Hintermann, Demain and Piret, Evidence for the presence of β-lactamase in *Streptomyces glaucescens* and its inhibition by sodium clavulanate. J. Indust. Microbiol. 16, 261 (1996). Demain, Klapatch, Jung and Lynd, Recombinant DNA technology in development of an economical conversion of waste to liquid fuel. Ann. NY Acad. Sci. 782, 402 (1996). Demain, Fungal secondary metabolism: regulation and functions. In: A Century of Mycology (B. Sutton, ed.) pp.233-253, Cambridge University Press, Cambridge (1996). Fang, Keables and Demain, Unexpected enhancement of β-lactam antibiotic formation in *Streptomyces clavuligerus* by very high concentrations of exogenous lysine. Appl. Microbiol. Biotechnol. 44, 705 (1996). Demain, Breaking through gridlock on the aromatic thruway (to lower cost specialty chemicals) Nature Biotechnol. 14, 580 (1996). Gift, Park, Paradis, Demain and Weaver, FACS-based isolation of slowly growing cells: double encapsulation of yeast in microdrops. Nature Biotechnol. 14, 884 (1996). Azuma and Demain, Interactions between gramicidin S and its producer, *Bacillus brevis*. J. Indust. Microbiol. 17, 56 (1996). Rius, Maeda and Demain, Induction of L-lysine ε-aminotransferase by L-lysine in *Streptomyces clavuligerus*, producer of cephalosporins. FEMS Microbiol. Lett. 144, 207 (1996). Demain and Fang, Secondary metabolism: The new biotechnology frontier. I n: Proceedings of Biotechnology Workshop (D. D.-Y. Ryu, S. S. Lee and M. H. Han, eds.), pp.1-12, Korean-American Scientists and Engineers Association, Rockville, MD. (1996). Fang, Pierson, Mishra, Koenig and Demain, Secondary metabolism in simulated microgravity: β-lactam production by *Streptomyces clavuligerus*. J. Indust. Microbiol. Biotechnol. 18, 22 (1997). Bykhovsky, Demain and Zaitseva, The crucial contribution of starved resting cells to the elucidation of the pathway of vitamin B12 biosynthesis. Crit. Revs. Biotechnol. 17, 21 (1997). Fang, Pierson, Mishra, Koenig and Demain, Gramicidin S production by *Bacillus brevis* in simulated microgravity. Curr. Microbiol. 34, 199 (1997). Romero, Martín, Liras, Demain and Rius. Partial purification, characterization and nitrogen regulation of the lysine ε-aminotransferase of Streptomyces clavuligerus. J. Indust. Microbiol. Biotechnol. 18, 241 (1997). Rius and Demain, Lysine ε-aminotransferase, the initial enzyme of cephalosporin biosynthesis in actinomycetes. J. Microbiol. Biotechnol. 7, 95 (1997). Fang and Demain, Influence of aeration and carbon source on production of microcin B17 by *Escherichia coli* ZK650. Appl. Microbiol. Biotechnol. 47, 547 (1997). Yashphe, Davis, Peng, Bok and Demain. New microorganism which convert compactin to pravastatin. Actinomycetologia 11, 20 (1997). Reynolds and Demain. Rapamycin, FK506, and ascomycin-related compounds. In: Biotechnology of Antibiotics, 2nd edition (W. R. Strohl, ed.), pp.497-520, Marcel Dekker, Inc., New York (1997). Fang, Pierson, Koening, Mishra and Demain, Effect of simulated microgravity and shear stress on microcin B17 production by Escherichia coli and on its excretion into the medium. Appl. Environ. Microbiol. 63, 4090 (1997). Lee, Kojima and Demain, Effect of nitrogen source on biosynthesis of rapmycin by Streptomyces hygroscopicus. J. Industr. Microbiol. Biotechnol. 19, 83 (1997). Rius & Demain, Regulation of lysine e-aminotransferase by carbon source and lack of control by phosphate in Streptomyces clavuligerus. Appl. Microbiol. Biotechnol. 48, 735-737 (1997). Martin, Gutierrez and Demain, β-Lactams. In: Fungal Biotechnology (T. Anke, ed.), pp. 91-127, Chapman & Hall, London (1997). Demain, Microbial natural products: Alive and well in 1998. Nature Biotechnology 16, 2-3 (1998). Lin, Lee, Hsu, Hamel and Demain, Properties of acetate kinase activity in *Clostridium thermocellum* cell extracts. Appl. Biochem. Biotechnol. 69, 137-145 (1998). A.L. Demain, H.J. Phaff and C.P. Kurtzman, The

industrial and agricultural significance of yeasts. In The Yeasts, A Taxonomic Study. 4th Ed. (C.P. Kurtzman and J.W. Fell, eds.), pp. 13-92, Elsevier, Amsterdam (1998). G.K. Wong, S. Griffith, I. Kojima and A.L. Demain, Antifungal activities of rapamycin and its derivatives prolylrapamycin, 32-desmethylrapamycin, and 32-desmethoxyrapamycin. J. Antibiot. 51: 487-491 (1998). K.N. Timmis and A.L. Demain, Ecology and industrial microbiology: strange bedfellows. Curr. Opin. Microbiol. 1: 267-270 (1998). Y. Peng and A.L. Demain, A new hydroxylase system in *Actinomadura* sp cells converting compactin to pravastatin. J. Indust. Microbiol. Biotechnol. 20: 373-375 (1998). I. Kojima and A. L. Demain, Preferential production of rapamycin *vs* prolylrapamycin by *Streptomyces hygroscopicus*. J. Indust. Microbiol. Biotechnol. 20: 309-316 (1998). H. Cho, J. L. Adrio , J.M. Luengo, S. Wolfe, S. Ocran, G. Hintermann, J. Piret and A.L.Demain, Elucidation of conditions allowing conversion of penicillin G and other penicillins to deacetoxycephalosporins by resting cells and extracts of *Streptmyces clavuligerus* NP1. Proc. Natl. Acad. Sci. USA 95: 11544-11548 (1998). G. Lancini and A. L. Demain, Secondary metabolism in bacteria: antibiotic pathways, regulation and function. In: Biology of the Prokaryotes. J. W. Lengeler, G. Drews and H. G. Schlegel, eds., pp. 627-651, Thieme, Stuggart (1999). G. Ozcengiz, J.-H. Kim, W. R. Lin, E. Ozcengiz, D. Westenberg, L. R. Lynd and A. L. Demain, Superiority of the PCR-based approach for cloning the acetate kinase gene of *Clostridium thermocellum*. J. Indust. Microbiol. Biotechnol. 21: 145-149 (1998). W. R. Lin, Y. Peng, S. Lew, C. C. Lee, J. J. Hsu, J.-F. Hamel and A. L. Demain, Purification and characterization of acetate kinase from *Clostridium thermocellum*. Tetrahedron 54: 15915-15925 (1998). A. L. Demain, J. F. Martin and R. P. Elander, Penicillin biochemistry and genetics. In: Penicillin: A Paradigm for Biotechnology. R. 1. Mateles, ed., pp. 93-114, Candida Corp., Chicago (1998). A. L. Demain and J. Zhang, Cephalosporin C production: the methionine story. Crit. Revs. Biotechnol. 18: 283-294 (1998). A. L. Demain, Induction of microbial secondary metabolism. Internatl Microbiol. 1: 259-264 (1998). A. L. Demain, Metabolites, primary and secondary. In:Encyclopedia. of Bioprocess Technology: Fermentation, Biocatalysis, and Bioseparation. M. C. Flickinger and Stephen W. Drew, eds., vol. 3, pp. 1713-1732, John Wiley and Sons, New York (1999). Y. Peng, E. A. Walker, J. C. Davis and A. L. Demain, A chemically-defined medium supporting growth and providing cells converting compactin to pravastatin. J. Indust. Microbiol. Biotechnol. 22: 78-79 (1999). A. L. Demain, Stunning achievements of industrial microbiology, ASM News 65: 311-316 (1999). A. L. Demain and R. P. Elander, The β-lactam antibiotics: past, present, and future. Antonie van Leeuwenhoek 75: 5-19 (1999). J. L. Adrio, H. Cho, J. M. Piret, A.. L. Demain, Inactivation of deacetoxycephalosporin C synthase in extracts of *Streptomyces clavuligerus* during bioconversion of penicillin G to deacetoxycephalosporin G, Enzyme and Microbial Technology 25: 497-501 (1999). M.J.Fernandez, J.L.Adrio, J.M.Piret, S.Wolfe, S.Ro and A.L.Demain, Stimulatory effect of growth in the presence of alcohols on biotransformation of penicclin G into cephalosporin-type antibiotics by resting cells of *Streptomyces clavuligerus* NP1, Appl. Microbiol. Biotechnol. 52: 484-488 (1999). A.L.Demain, Pharmaceutically active secondary metabolites of microorganisms, Appl. Microbiol. Biotechnol. 52: 455-463 (1999). M.A.Baez-Vasquez, J.L.Adrio, J.M.Piret and A.L.Demain, Further studies on the bioconversion of Penicillin G into deacetoxycephalosporin G by resting cells of *Streptomyces clavuligerus* NP-1, Appl. Biochem. Biotechnol. 81: 145-152 (1999). A.L. Demain, Microbial biotechnology, Trends Biotechnol. 18: 26-31 (2000). A.Fang, G.K.Wong & A.L.Demain, Enhancement of the antifungal activity of Rapamycin by the coproduced Elaiophylin and Nigericin, J.of Antiobiotics. 53: 158-162 (2000). A. Fang, D. L. Pierson, S. K. Mishra and A. L. Demain, Growth of *Streptomces hygroscopicus* in rotating-wall bioreactor under simulated microgravity inhibits rapamycin production, Appl. Microbiol. Biotechnol. 54: 33-36 (2000). A. Fang, D. L. Pierson, S. K. Mishra and A. L. Demain, Relief from glucose interference in microcin B17 biosynthesis by growth in a rotating-wall bioreactor, Lett. Appl. Microbiol. 31: 39-41(2000). M. A. Baez-Vasquez and A. L. Demain, Immobilized *Streptomyces clavuligerus* NP1 cells for biotransformation of penicillin G into deacetoxycephalosporin G, Appl. Biochem. Biotechnol. 87: 135-140 (2000). W-S. Kim, Y. Wang, A. Fang and A. L. Demain, Methionine interference in rapamycin production involves repression of demethylrapamycin methyltransferase and S-adenosylmethionine synthetase. Antimicrob. Agents Chemother. 44: 2908-2910 (2000). Y. Peng and A. L. Demain, Bioconversion of compactin to pravastatin by *Actinomadura* sp. ATCC 55678. J. Molec. Catalysis B: Enzymatic 10: 151-156 (2000). A. L. Demain, Idiolites from microorganisms: a brilliant past, a bright future. In: 2000 Years of Natural Products Research-Past, Present and Future. T. J. C. Luijendijk ed., pp. 15-39, Phytoconsult, Leiden (2000). A. L. Demain, Microbial natural products: a past with a future. In: Biodiversity- New Leads for the Pharmaceutical and Agrochemical Industries. S. K. Wrigley, M. A. Hayes, R. Thomas, E. J. T. Chrystal and N. Nicholson, eds., pp. 3-16, Royal Society of Chemistry, Cambridge (2000). A. L. Demain, Small bugs, big business: The economic power of the microbe. Biotechnol. Adv. 18: 499-514 (2000). A. L. Demain and A. Fang, The natural functions of secondary metabolites. In: History of

Modern Biotechnology I, A. Fiechter, ed., pp. 1-39, Springer, Berlin (2000). A. L. Demain, The changing face of industrial microbiology/biotechnology. In: Proceedings of the Symposium on New Directions in Biotechnology, pp. 1-6, Society of Biotechnology and Bioengineering of Japan, Osaka University (2000). A. L. Demain, Edward P. Abraham, Cell-free systems and the fungal biosynthesis of beta-lactams. J. Antibiotics. 53: 995-1002 (2000). A. L. Demain and G. Lancini, Bacterial pharmaceutical products. In: Prokaryotes, An Evolving Electronic Resource for the Microbiological Community. M. Dworkin, ed-in-chief, Springer-Verlag, New York (2001). W. H. Strohl, H. B. Woodruff, R. L. Monaghan, D. Hendlim, S. Mochales, A. L. Demain and J. Liesch, The history of natural products research at Merck & Co., Inc. SIM News. 51 (1): 5-20 (2001). A. L. Demain, J. L. Adrio and J. M. Piret, Bioconversion of penicillins to cephalosporins. In: Enzyme Technologies for Pharmaceutical and Biotechnological Applications, H. A. Kirst, W-K Yeh and M. J. Zmijewski, Jr, eds., pp. 61-88, Marcel Dekker, Inc. New York (2001). A. L. Demain and A. Fang, Secondary metabolism in simulated microgravity. Chem Rec 1: 333-346 (2001). Q. Gao, A. Fang, D. L. Pierson, S. K. Mishra and A. L. Demain, Shear stress enhances microcin B17 production in a rotating wall bioreactor, but ethanol stress does not. Appl Microbiol Biotechnol 56: 384-387 (2001). Q. Gao, A. Fang. and A. L. Demain, Induction of microcin B17 formation in *Escherichia coli* ZK650 by limitation of oxygen and glucose is independent of glucose consumption rate. J Indust Microbiol & Biotechnol 26: 341-344 (2001). A. L. Demain, Molecular genetics and industrial microbiology-30 years of marriage. J. Indust. Microbiol. Biotechnol. 27: 352-356 (2001).Q. Gao and A. L. Demain, Improvement in the bioconversion of penicillin G to deacetoxycephalosporin G by elimination of agitation and addition of decane. Appl. Microbiol. Biotechnol. 57: 511-513 (2001). Q. Gao and A. L. Demain, Effect of solvents on bioconversion of penicillin G to deacetoxycephalosporin G. J. Antibiot. 54: 958-961 (2001). A. L. Demain, Relationship between universities and industry: the American university perspective. Food Technol. Biotechnol. 39: 157-160 (2001). Y. R. Cheng, J. Huang, H. Qiang, W. L. Lin, and A. L. Demain, Mutagenesis of the rapamycin producer *Streptomyces hygroscopicus* FC904. J. Antibiot. 54: 967-972 (2001). A. L. Demain, Biotechnology: an environment where microbiology, genetics and engineering meet. In: Current Studies of Biotechnology. II. Environment, Z. Kniewald, ed., pp. 1-11, Croatian Society of Biotechnology, Zagreb (2001). W.-S. Kim, L. Xu, D. Souw, A. Fang, and A. L. Demain, An unexpected inhibitory effect of rapamycin against germination of spores of *Bacillus brevis* strain Nagano. J Antibiot. 55: 650-654 (2002). M. Kitabatake, K. Ali, A. Demain, K. Sakamoto, S. Yokoyama, and D. Söll, Indolmycin resistance of *Streptomyces coelicolor* A3(2) by induced expression of one of its two tryptophanyl-tRNA synthetases. J. Biol. Chem. 277: 23882-23887 (2002). J. L. Adrio and A. L. Demain, Improvements in the formation of cephalosporins from penicillin G and other penicillins by bioconversion. Org. Proc. Res. & Devel. 6:427-433 (2002). J. L. Adrio, G. A. Hintermann, A.L. Demain and J. M. Piret, Construction of hybrid bacterial deacetoxycephalosporin C synthases (expandases) by in vivo homeologous recombination. Enzyme Microb. Technol. 31:932-940 (2002). S. Sanchez and A. L. Demain, Metabolic regulation of fermentation processes. Enzyme Microb. Technol. 31:895-906 (2002). J.F. Martin and A.L. Demain. Unraveling the methionine-cephalosporin puzzle in *Acremonium chrysogenum*. Trends Biotech. 20:502-507 (2002). A.L. Demain and P. Vaishnav, Regulation of β-lactam antibiotic biosynthesis by carbon sources. Chimica Oggi/Chemistry Today 20 (11-12):46-51 (2002). A. L. Demain and R. F. Wolfe, Marvin P. Bryant 1925-2000, A biographical memoir. In: Biographical Memoirs 81, pp. 3-34. The National Academy Press, Washington D.C. (2002). A. L. Demain, Prescription for an ailing pharmaceutical industry. Nature Biotechnol 20: 331 only (2002). Q. Gao and A. L. Demain, Improvement in the resting-cell bioconversion of penicillin G to deacetoxycephalosporin G by addition of catalase. Lett. Appl. Microbiol. 34: 290-292 (2002). W. Yang, E.A. Hartwieg, A. Fang and A.L. Demain. Effects of carboxymethylcellulose and carboxypolymethylene on morphology of *Aspergillus fumigatus* NRRL 2346 and fumagillin production. Curr. Microbiol. 46: 24-27 (2003). Q. Gao, J.M. Piret, J.L. Adrio and A.L. Demain. Performance of a recombinant strain of *Streptomyces lividans* for bioconversion of penicillin G to deacetoxycephalosporin G, J. Indust. Microbiol. 30: 190-194 (2003).W. Yang, W.-S. Kim, A. Fang and A.L. Demain. Carbon and nitrogen source nutrition of fumagillin biosynthesis by *Aspergillus fumigatus*. Curr. Microbiol. 46: 275-279 (2003).W.-S. Kim, S. Davis, G. Wong and A.L. Demain. Nutritional studies on the growth of the rapamycin-producing *Streptomyces hygroscopicus*. J. Microbiol. Biotechnol. 13: 560-563 (2003). A.L. Demain. Herman Jan Phaff: professor, mentor, friend and colleague. Int. Microbiol. 6: 157-161 (2003). J.L. Adrio and A.L. Demain. Fungal biotechnology. Int. Microbiol. 6: 191-199 (2003). A. L. Demain and P. Vaishnav. Nitrogen regulation of biosynthesis of antibiotics and other secondary metabolic products. PharmaChem 2 (No. 10): 92-96 (2003). A. L. Demain and P. Vaishnav. Secondary metabolism in microbes and its control by phosphate and metals. SIM News 54: 104-113 (2004). A. L. Demain. The biopharmaceutical revolution. Chimica Oggi/Chemistry Today. 22 (Nos. 11-12): 41-44

(2004). A. L. Demain, J. Velasco and J. L. Adrio. Industrial mycology: past, present, and future. In: Handbook of Industrial Mycology (Z. An, Ed.) pp. 1-25. Marcell Dekker, New York (2004). A. L. Demain. Pickles, pectin and penicillin. Annu. Rev. Microbiol. 58:1-42 (2004). A. L. Demain and Ajai Sehgal. In memoriam: Surendra Nath Sehgal, the godfather of rapamycin (1932-2003). SIM News 54:20-21 (2004). A.L. Demain, M. Newcomb and J.H. David Wu. Cellulase, clostridia and ethanol. Microbiol. Molec. Biol. Rev. 69: 124-154 (2005). J. L. Adrio and A. L. Demain. Microbial cells and enzymes; a century of progress. In: Methods in Biotechnology, vol. 17 (J. L. Barredo, Ed.) pp. 1-27. Humana Press, Totowa, New Jersey (2005). A. L. Demain, D. F. Gerson and A. Fang. Effective levels of tetanus toxin can be made in a production medium totally lacking both animal (e.g., brain heart infusion) and dairy proteins or digests (e.g., casein hydrolysates). Vaccine 23: 5420-5423 (2005). A. L. Demain and Lixin Zhang. Natural products and drug discovery. In: Natural Products: Drug Discovery and Therapeutic Medicine. (L. Zhang and A. L. Demain, Eds.) pp. 3-29. Humana Press, Totowa, New Jersey (2005).

Ferm - organic acids - pyruvic   DemAri

Available online at www.sciencedirect.com

SCIENCE DIRECT

ELSEVIER

Enzyme and Microbial Technology 36 (2005) 832–839

ENZYME and MICROBIAL TECHNOLOGY

www.elsevier.com/locate/enzmictec

Metabolic engineering of *Torulopsis glabrata* for improved pyruvate production

Qinhong Wang, Peng He, Dajun Lu, An Shen, Ning Jiang*

Center for Microbial Biotechnology, Institute of Microbiology, Chinese Academy of Sciences, P.O. Box 2714, 13 Bei 1 Tiao, Zhongguancun, Beijing 100080, P.R. China

Received 13 April 2004; accepted 10 January 2005

Abstract

During pyruvate production, ethanol is produced as a by-product, which both decreases the amount of pyruvate and makes the recovery of pyruvate more difficult. Pyruvate decarboxylase (PDC, EC 4.1.1.1), which degrades pyruvate to acetaldehyde and ultimately to ethanol, is a key enzyme in the pyruvate metabolism of yeast. Therefore, to order to increase the yield of pyruvate in *Torulopsis glabrata*, targeted *PDC*-disrupted strains were metabolically engineered. First, *T. glabrata ura3* strains that were suitable for genetic transformation were isolated and identified through ethyl methansulfonate mutagenesis, 5-fluoroortic acid media selection, and *Sacchramyces cerevisiae URA3* complement. Next, the *PDC* gene in *T. glabrata* was specifically disrupted through homologous recombinant with the *S. cerevisiae URA3* gene as the selective marker. The PDC activity of the disruptants was about 33% that of the parent strain. Targeted *PDC* gene disruption in *T. glabrata* was also confirmed by PCR amplification and sequencing of the *PDC* gene and its mutants, PDC activity staining, and PDC Western blot. The disruptants displayed higher pyruvate accumulation and less ethanol production. Under basal fermentation conditions (see Section 2), the disruptants accumulated about 20 g/L of pyruvate with 4.6 g/L of ethanol, whereas the parental strain (*T. glabrata* IFO005) only accumulated 7–8 g/L of pyruvate with 7.4 g/L of ethanol. Under favorable conditions in jar fermentation, the disruptants accumulated 82.2 g/L of pyruvate in 52 h.

Keywords: Pyruvate; Pyruvate decarboxylase; *ura3* Auxotroph; Gene disruption; *Torulopsis glabrata*; Ethanol; Metabolic engineering

1. Introduction

Pyruvate is a key metabolite in the glycolytic pathway; it is also a model product for studying the regulation of glycolysis [1]. Pyruvate is the key substrate for the enzymatic production of amino acids such as L-tryptophan, L-tyrosine, and L-dihydroxyphenylalanine (L-DOPA) [2,3]. The commercial demand for pyruvate has been increasing due to its use as a starting material for the synthesis of many drugs and agrochemicals, and as a component of animal cell cultures. Pyruvate also has a healthcare function (i.e., it has been recently approved as a dietary supplement) [4].

Pyruvate is associated with many metabolic pathways in microorganisms and it is not usually accumulated or secreted at any significant level by microorganisms. However, under certain conditions excess pyruvate production has been observed from bacteria [5], yeasts [6], molds, and basidiomycetes [7]. With the advent of various modern molecular biotechniques, many higher-pyruvate-producing strains were successfully engineered [8–10]. Optimum process control strategies have also made great contributions to enhancing pyruvate production [11,12].

In order to get a high yield and increase the productivity, it is necessary to accelerate glycolysis and/or to inhibit the activities of the enzymes that are responsible for pyruvate degradation in yeasts [13]. Generally, three key enzymes-pyruvate decarboxylase (PDC), pyruvate dehydrogenase complex (PDH complex), and pyruvate carboxylase (PYC) are responsible for pyruvate catabolism in yeasts [1]. These enzymes all need vitamins or their derivatives as the coenzymes. Improving cultivation conditions, especially the concentrations

* Corresponding author. Tel.: +86 10 62553081; fax: +86 10 62560912.
E-mail address: jiangn@sun.im.ac.cn (N. Jiang).

0141-0229/$ – see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.enzmictec.2005.01.015

of required vitamins, was a possible way to control enzyme activity and direct the metabolism of yeast to accumulate high levels of pyruvate [6,14]. Using vitamin limitation as a strategy for metabolic control would be efficient but not perfect for pyruvate production because it would require many tricks; this is due to feeding different species and concentrations of vitamins in the process of fermentation. Other process strategies, such as oxygen supply control [11], also have some tricks. Treating wild-type pyruvate-producing yeast, *T. glabrata* IFO005, with chemical mutagens can generate mutants with reduced activity of pyruvate decarboxylase; thus increasing pyruvate production and decreasing ethanol production [15]. However, the genetic and metabolic profiles of mutants derived from chemical mutagenesis were poorly characterized and mutagenesis remained a random process where science was complemented with elements of art.

The aim of metabolic engineering is defined as the purposeful modification of metabolic networks in living cells to produce desirable chemicals with superior yields and productivity by using recombinant DNA technologies [16]. First suggested by Dr. Bailey in 1991, metabolic engineering powered by techniques from applied molecular biology and reaction engineering has since become very popular in the last decade. Major efforts have been made in selecting the appropriate microorganisms to both efficiently and abundantly produce chemicals of commercial interest. The notion of metabolic engineering significantly extends the range of microbes utilizable as productive hosts. In order to achieve these objectives, three main research approaches are usually employed: (1) introducing exogenous genes which convert the final precursor of a host organism to a desirable chemical at a viable yield; (2) enhancing metabolic flux through a pathway to increase the synthesis of the final precursor or product; and (3) minimizing the biosyntheses of by-products (or other products). There have been great changes in strain improvement since the dictionary of metabolic engineering.

In this particular, the third research approach of metabolic engineering (3) was applied to specifically disrupt the *PDC* gene in a pyruvate-producing yeast *T. glabrata*. As a result, the carbon flux was redirected in the yeast strain from ethanol formation with respect to increased production of pyruvate. Also, this study shows how PDC influences the pyruvate yield in *T. glabrat*. The resultant strains with higher pyruvate productivities and commercial potentials were also described and discussed.

2. Material and methods

2.1. Strains, plasmids, and culture conditions

Table 1 lists each strain and plasmid used for this study. *Escherichia coli* DH5α was used for plasmid construction, and *T. glabrata* IFO005 (a gift from Dr. Liu Jiquan, South Yangtze University, P.R. China) was used as the recipient strain. Recombinant *E. coli* strains were grown at 37 °C in Luria-Bertani medium [17] supplemented with ampicillin (100 mg/L). *T. glabrata* was grown in YEPD (10 g yeast extract, 20 g tryptone, 20 g glucose, pH 5.5) at 30 °C, 5-fluoroorotic acid (5-FOA) medium [18] or synthetic medium (SD) [19] except for pyruvate production.

Table 1
Stains and plasmids

Strains or plasmid	Related characteristic	Source or reference
Strains		
E. coli DH5α	*supE44ΔlacU169(φ80lacZΔM15) hsdR17recA1endA1gyrA96thi-1relA1*	Stratagene
T. glabrata		
IFO005	Wild type	Dr. Liu Jiquan
IFO005-3	*ura3*	This work
IFO005-7	*ura3*	This work
IFO005-36	*ura3 pdc::URA3*	This work
IFO005-72	*ura3 pdc::URA3*	This work
Plasmids		
pET21a(+)		Navogen
pRUL129		Dr. H. Yde Steensma
pETP1		This work
pEPU1		This work

2.2. Media for pyruvate production and its fermentation conditions

The seed medium with a pH of 5.5 was composed of 10% glucose, 3% fish peptone, 0.1% KH_2PO_4, 0.05% $MgSO_4{\cdot}7H_2O$, 4 mg/L of nicotinic acid (NA), 200 μg/L of pyridoxine, 2 μg/L of biotin and 4% $CaCO_3$. The basal fermentation medium with a pH of 5.5 was composed of 10% glucose, 0.1% soybean peptone, 0.6% $(NH_4)_2SO_4$, 0.1% KH_2PO_4, 0.05% $MgSO_4{\cdot}7H_2O$, 8 mg/L of NA, 1 mg/L of pyridoxine, 30 μg/L of biotin, 30 μg/L of thiamine and 4% $CaCO_3$. The yeast was cultivated in Erlenmeyer flasks at 30 °C at 200 rpm/min for 24–36 h. The jar fermentation medium with a pH of 5.5 was composed of 15% glucose, 3% fish peptone, 0.1% KH_2PO_4, 0.05% $MgSO_4{\cdot}7H_2O$, 4 mg/L of NA, 0.1 mg/L of pyridoxine, 10 (g/L of biotin, 10 μg/L of thiamine, pH was maintained by supplementing of 5 N NaOH, agitation was 500 rpm and air supply was 1:1. The inoculum was 10% of the total volume of the medium.

2.3. Isolation and characterization of T. glabrata ura3 strains

Ethyl methansulfonate (EMS, Sigma Chemical Co., St. Louis, Mo.) mutagenesis of *T. glabrata* was performed as described for *Saccharomyces cerevisiae* [19]. Unless stated otherwise, EMS treatment killed approximately 50% of the initial population. *ura* (uracil auxotrophic phenotype) mutants were obtained on 5-FOA agars (*ura* mutants can grow on 5-FOA mediums, whereas others cannot). Plasmid pRUL129

Table 2
Oligonucleotide primers used in this study

Oligonucleotide primer	DNA sequence
PDC-C11	5′-GCGGATCCATGTCTGAAATTACTTTGG-3′
PDC-C12	5′-GTTCTCGAGCTTATTGCTTAGCGTTGG-3′
ScURA3-C1	5′-AGCAAGCTTTGAGAGTGCACCATAC-3′
ScURA3-C2	5′-GTTGGTACCTCTCCTTACGCATCTG-3′
TgURA3-C1	5′-CTTACAATGTCCAGTGCC-3′
TgURA3-C2	5′-TTGGATATGTCTTGAATCAG-3′
PDC-C21	5′-ATGTCTGAAATTACTTTGGT-3′
PDC-C22	5′-TTATTGCTTAGCGTTGGTGGAA-3′
PDC-C31	5′-GAGACCAGACTAATACAACT-3′
PDC-C32	Same as PDC-C12

[20] was transformed into *ura* mutants by modified lithium acetate method [19], and *ura3* mutants that could be complemented by the *S. cerevisiae URA3* gene on the plasmid pRUL129 were confirmed.

2.4. Construction plasmids and strains

The primers used in this study are shown in Table 2. The *T. glabrata PDC* gene (accession number AF545432) was amplified using primers PDC-C11 (forward) and PDC-C12 (reverse) with the genomic DNA of *T. glabrata* as the template so that a 1710-bp DNA fragment (Fragment I, Fig. 1A) containing a 5′*Bam*HI and 3′ *Xho*I site (underlining in the primers) was obtained. The genomic DNA of *T. glabrata* was prepared by the glass-beads method [19]. Fragment I was digested with *Bam*HI and *Xho*I to obtain an about 1.7-kb DNA fragment containing *T. glabrata PDC* coding region. This DNA fragment was then cloned directly into pET21a(+) to produce plasmid pETP1. The *S. cerevisiae URA3* gene was amplified using primers URA3-C1 (forward) and URA3-C2 (reverse) with plasmid pRUL129 as a template so that a 1182-bp DNA fragment (Fragment II, Fig. 1A) containing a 5′ *Hin*dIII and 3′ *Kpn*I site (underlining in the primers) was obtained. Plasmid pETP1 and Fragment II were digested with *Hin*dIII and *Kpn*I to obtain a 6470 and 1176-bp fragment, respectively. These two DNA fragments was ligated with T4 DNA ligase to produce plasmid pEPU1, which contained an about 2250-bp DNA fragment (Fragment III, Fig. 1A) including *T. glabrata PDC-F1*, *S. cerevisiae URA3*, and *T. glabrata PDC-F2*. *PDC-F1* and *PDC-F2* were located in the region of nucleotide 1–546 and 1170–1692, respectively, in *T. glabrata PDC* coding region. Fragment III was amplified using primers PDC-C21 (forward) and PDC-C22 (reverse) with plasmid pEPU1 as a template, and was transformed into *T. glabrata ura3* stains (IFO005-3 and IFO005-7) by the modified lithium acetate method. In vivo, fragment III replaced *PDC* gene coding region (1,692-bp) of *T. glabrata* IFO005-3 and IFO005-7 by homologous recombination, and *PDC* gene-disrupted *T. glabrata* strains IFO005-36 and IFO005-72 were constructed.

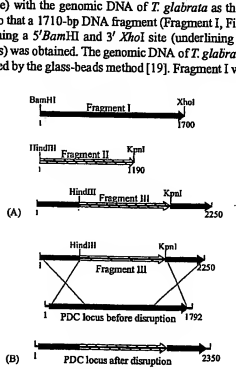


Fig. 1. Schematic representation of the DNA fragments for gene disruption and the disruption of *PDC* by homologous recombinant. *T. galbrata PDC* or fragments (black), *S. cerevisiae URA3* gene (*ScURA3*, hatched). (A) Structure of DNA fragments I–III. Fragment I denoted PCR product of *T. galbrata PDC*; Fragment II denoted PCR product of *S. cerevisiae URA3*; Fragment III denoted PCR product that containing *T. glabrata PDC-F1*, *S. cerevisiae URA3* and *T. glabrata PDC-F2* (from left to right); (B) *PDC* gene disruption by homologous recombination.

2.5. Analytical procedures

The fermentation broth was centrifuged at 10,000 × *g* for 20 min and the supernatant was used to determine metabolites. Pyruvate in the culture medium was determined by the DNP (2, 4-dinitrophenylhydrazine) method [21] with sodium pyruvate as standard. Glucose was determined by the DNS (3, 5-dinitrosalicylate) method [22]. Glycerol was determined by the Nash method [23]. Ethanol was determined by HPLC with a refractive index (RI) detector and an Aminex HPX-87H column (Bio Rad Laboratories, USA); the mobile phase was a 5 mM H_2SO_4 aqueous solution with a flow rate of 0.6 mL/min; the injection volume was 25 μL. Pyruvate, glycerol, and glucose concentration were also determined by HPLC under the same conditions. Yeast cell growth (biomass) was measured turbidimetrically at 660 nm (OD_{660}) after the culture broths were diluted 50-fold with water. The optical density value was converted to dry cell weight (DCW) using the calibration equation (1 OD_{660} = 0.3 g DCW/L).

2.6. Preparation of cell extracts and enzyme assays

T. glabrata IFO005 or its mutants were cultured aerobically at 30 °C for 16 h in 18 mm-ϕ tube containing 5 mL of YEPD or YEPD supplemented with 20 mg/L of uracil. Cells collected from the two media were washed with 20 mM sodium phosphate buffer (pH 6.0) and re-suspended in the same buffer. After being vertoxed with glass beads for 10 min at 4 °C, the cell extracts were centrifuged at 15,000 × *g* for 30 min to remove the glass beads and cell debris; PDC ac-

tivity in the supernatant was measured by a coupled enzyme assay [24] with a minor modification. Assays were performed at 30 °C and pH 6.0 in 100 mM citrate buffer, 33 mM sodium pyruvate with 1.5 units of yeast alcohol dehydrogenase (ADH) in 3 mL volume, unless otherwise noted. Enzyme activity was measured for the first 1–3 min.

2.7. Enzyme activity staining and Western blot analysis

Cell extracts were subjected to the native gradient PAGE using a 5–12.5% linear gradient slab gel with a running buffer consisting of 25 mM Tris and 192 mM glycine (pH 8.8) at 4 °C. Activity staining with 1,2-dianilinoethane [25] was utilized for the native gel. The protein extracts prepared for the enzyme activity assays were used in the Western blot analysis. Western blotting was performed as described [26] with some modifications and stained with dissolved diaminobenzidine (50 mL phosphate buffer containing 30 mg DAB and 75 μL 30% H_2O_2). PDC was probed with a rabbit antiserum directed against the *T. glabrata* purified PDC protein (unpublished data).

2.8. Other methods

Protein concentration was determined with the Bradford method [27] using bovine serum albumin (Roche Molecular Biochemicals) as the standard. Basic molecular biology techniques were performed according to the manufactures' instructions or standard procedures [17]. SDS-polyacrylamide gel electrophoresis (SDS-PAGE) with SDS-Tris system using 10% polyacrylamide gel was carried out according to the procedure described by Laemmli [28] using the Mini-Protean III apparatus (Bio-Rad Laboratories, Inc.).

3. Results

3.1. Isolation and characterization of ura3 mutants of T. glabrata

To generate the *T. glabrata* strain suitable for genetic transformation, strains bearing mutation in the gene (*URA3*) encoding orotidine 5′-phosphate decarboxylase were isolated. The *T. glabrata* IFO005 was used as the parental wild type strain. Stable *ura* mutants were identified after EMS mutagenesis [19] and 5-FOA media selection [18], and purified for single colonies. Since the *ura* phenotypes of *T. glabrata ura* mutant strains could be complemented by the *S. cerevisiae URA3* gene on the episomal plasmid pRUL129, the two *ura3* stains designated as IFO005-3 and IFO005-7 were confirmed. The growth of the parental strain (*T. glabrata* IFO005), *ura3* mutants (*T. glabrata* IFO005-3 and IFO005-7), and *S. cerevisiae URA3* complemented mutants (*T. glabrata* IFO005-36 and IFO005-72) in SD, SD supplemented uracil (20 mg/L), and 5-FOA agars was shown in Fig. 2, all of which gave the evidences of our experiment about screening mutants. Due

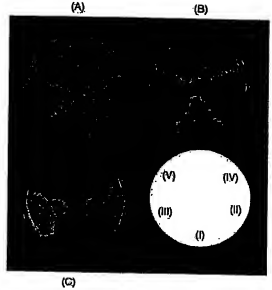


Fig. 2. The growth of *T. glabrata* in SD supplement uracil (200 mg/l) (A), SD (B), and 5-FOA agars (C). (I) *T. glabrata* IFO005; (II) *T. glabrata* IFO005-3; (III) *T. glabrata* IFO005-7; (IV) *T. glabrata* IFO005-36 and (V) *T. glabrata* IFO005-72.

to the *URA3* mutation, strains IFO005-3 and IFO005-7 could grow on 5-FOA agar, but not on SD agar. However, strains IFO005-36 and IFO005-72 complemented by *S. cerevisiae URA3* could grow on SD agar, but not on 5-FOA agar. The parental strain IFO005 could also grow on SD agars, but not on 5-FOA agar because of its normal *UAR3* gene. All strains could grow on SD supplement uracil agar.

By the PCR cloning and sequencing of the *URA3* genes of *T. glabrata* IFO005-3 and IFO005-7 with primers TgURA3-1 and TgURA3-2 (Table 2), we found that some amino acid residues in the *URA3* gene in these two mutants were mutative; two mutations, E185K and T191R, were produced in both mutants. IFO005-3 had the other mutation, C33F and IFO005-7 possessed the other four mutations, K29T, G108R, G128R and P202H. However, all of these mutated amino acids were not conserved [29,30]. We could not tell what gave rise to the *ura3* auxotroph of *T. glabrata* IFO005-3 and IFO005-7, but these mutations should be relevant to *ura3* auxotroph.

3.2. Construction of pdc mutants of T. glabrata

To achieve increased pyruvate production and reduced ethanol production, the most promising and efficient approach was to minimize metabolic flux to the pathways of by-products (ethanol) by metabolic engineering. We disrupt the *PDC* gene in *T. glabrata* to block the pathway from pyruvate to aldehyde, then to ethanol, thus increasing pyruvate accumulation and decreasing ethanol production.

During the construction of *pdc* strains of *T. glabrata*, *PDC* gene disruption was carried out by homologous recombination (Fig. 1B). The constructed homologous DNA fragment

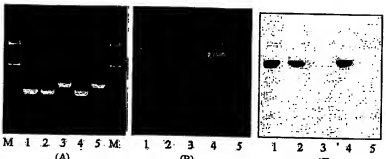


Fig. 3. Characterization of *T. glabrata*. (A) PCR analysis of *T. glabrata*; (B) Pdc activity of crude extract from *T. glabrata*; (C) Western Blot analysis of crude extract from *T.glabrata*. (1) *T. glabrata* IFO005; (2) *T. glabrata* IFO005-3; (3) *T. glabrata* IFO005-36; (4) *T. glabrata* IFO005-7; (5) *T. glabrata* IFO005-72; M) DNA maker (λ DNA/*EcoRI*+*HindIII*: 21,227, 5148, 4973, 4268, 3530, 2027, 1904, 1584, 1375, 947, 831, 564 and 125 bp).

(Fragment III) was introduced into *T. glabrata ura3* strains (IFO005-3 and IFO005-7) by lithium acetate transformation, and disruptants appeared after 2–3 days on agar SD medium. *pdc* strains of *T. glabrata* IFO005-36 and IFO005-72 were obtained from IFO005-3 and IFO005-7, respectively.

Based on our laboratory experience, one-step gene disruption in the yeast *T. glabrata* was not straightforward. After a systematic study of several factors that could influence frequency of gene disruption, we found that the length of the target gene region flanking the marker gene is a critical factor in one-step gene disruption in *T. glabrata*. Targeted gene regions of about 500 bp flanking the marker were necessary to obtain a disruption frequency. Targeted gene disruption in some non-conventional yeasts has similar results [31].

3.3. Characterization of pdc strains of T. glabrata

Enzymatic analysis of *pdc* strains, *T. glabrata* IFO005-36 and IFO005-72, grown in YEPD medium demonstrated the presence of very low PDC activity (0.48 and 0.51 U/mg, respectively). Also PDC activities in *T. glabrata* IFO005, IFO005-3, and IFO005-7 were about 1.5–1.7 U/mg. These results indicated that *PDC* gene disruption significantly reduced the enzyme activities of the *pdc* strains. The total DNA isolated from five strains of IFO005 and its derivatives was subjected to PCR analysis using two primers PDC-C31 (forward) and PDC-C32 (reverse) (Table 2). A 2.25-kb fragment was detected in the two disruptants IFO005-36 and 005-72, and a 1.9-kb fragment was detected in wild type strain IFO005, and *ura3* mutants IFO005-3 and 005-7 (Fig. 3A); this was consistent with the disruption of *PDC* as predicted from the yeast genome sequence. Sequencing of their PCR products also detected that the *PDC* gene in parent strains was replaced by homologous DNA fragment (Fragment III), which contained *T. glabrata PDC-F1*, *S. cerevisiae URA3*, and *T. glabrata PDC-F2*. The primer PDC-C31 was in the upstream of PDC starting codon. PDC–activity staining of cell extracts from *T. glabrata* IFO005 and its mutants (Fig. 3B), and Western blot analysis (Fig. 3C) further confirmed *PDC* gene disruption in these mutants. These results showed that PDC-activity and immunologically cross-reacting materials were not detected in the cell extracts from IFO005-36 and IFO005-72. Therefore, *PDC* has been efficiently disrupted in these strains as compared to their parent strains, IFO005-3 and IFO005-7.

3.4. Growth and fermentation pattern of T. glabrata mutants and wild-type strains

Fig. 4 represents the growth on SD medium containing uracil (200 mg/mL) and YEPD medium of IFO005 *pdc* strains with respect to the wild-type and *ura3* strains. From

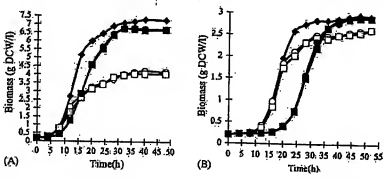


Fig. 4. Consequences *PDC* and *URA3* disruption for growth. (A) YEPD medium; (B) SD medium. (♦) IFO005; (○) IFO005-3;(●) IFO005-36; (□) IFO005-7; (■) IFO005-72.

Table 3
Pyruvate production by *T. glabrata* IFO005 and its derivatives under basal fermentation conditions

Strains	Biomass (g DCW/L)	Pyruvate (g/L)	Residual glucose (g/L)	Ethanol (g/L)
IFO005	18.21 ± 0.66	7.3 ± 0.5	7.6 ± 0.9	7.4 ± 0.4
IFO005-3	7.95 ± 0.39	4.7 ± 0.4	13.6 ± 0.9	7.1 ± 0.3
IFO005-7	7.56 ± 0.36	5.3 ± 0.5	13.1 ± 1.4	6.6 ± 0.4
IFO005-36	16.29 ± 0.36	19.9 ± 0.9	11.3 ± 1.0	4.6 ± 0.4
IFO005-72	15.90 ± 0.51	20.4 ± 1.5	10.2 ± 1.4	4.6 ± 0.4

Scale: 15 mL basal fermentation medium in a 500 mL Erlenmeyer flask, cultivation time (h): 30 h. Data are presented as the average (±) standard deviation for duplicate assays on two independent cultures.

the growth in two media, we found that *pdc* strains have a delay at the early stages of growth, especially in SD medium (Fig. 4B), which showed that *PDC* gene disruption really has an effect on cell growth. The metabolic process of pyruvate by PDC could contribute to $NADH/NAD^+$ balance [32]. Thus, *PDC* gene disruption would destroy the balance of $NADH/NAD^+$, which results in the delay of the cell growth of *pdc* mutants. Once $NADH/NAD^+$ balanced again through the regulation of other metabolic pathway, the cells also restored the growth similar to the wild type strain. As for *ura3* mutants, all growth in various basic media was affected due to the uracil auxotroph (Fig. 4A and B, Table 3).

The afore mentioned *pdc* mutants and their parental strains were isolated from single colonies, and subsequently tested for their ability to produce pyruvate in the basal fermentation medium. The data in Table 3 shows the advantages of *pdc* strains IFO005-36 and IFO005-72 over strains IFO005, IFO005-3, and IFO005-7 in pyruvate production: the former have higher pyruvate production and lower ethanol production in basal fermentation medium. Although IFO005 could grow faster than the *pdc* mutants, its high-biomass did not result in high production of pyruvate. Our previous paper [33] also showed that it was necessary to keep a balance between biomass and pyruvate production during fermentation. Under optimization conditions in 5 L-jar fermentation, the disruptant IFO005-36 accumulated the highest yield of 82.2 g/L of pyruvate in 52 h (Fig. 5), and the yield of pyruvate to glucose was 0.548. As the same time, it only accumulated about 4.7 g/L of ethanol.

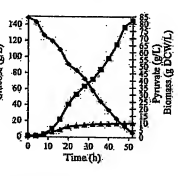


Fig. 5. Pyruvate fermentation by *T. glabrata* IFO005-36 in 5 L-jar fermentor. (■) Pyruvate; (♦) glucose and (▲) biomass.

4. Discussion

T. glabrata is a superior species for pyruvate and optically active α-hydroxyketone production [6,34]. In this investigation, we used EMS mutagenesis and complemented integration to screen the *ura3* auxotrophs. These *ura3* strains of *T. glabrata* should be useful in the metabolic engineering of *T. glabrata* for improved pyruvate or α-hydroxyketone production. *T. glabrata* could also be transformed with circular palsmid carrying ARS elements from *S. cerevisiae* [35] or from other *T. glabrata* strains [36], and the plasmid replicate as unstable, high-copy extrachromosomal element. This was also useful in the metabolic engineering of *T. glabrata* for improved pyruvate or α-hydroxyketone production. The 5-FOA selections were extremely useful for a number of genetic and molecular biological manipulations that required the detection of rare *ura3* cells. The advantages of the 5-FOA selection were the availability of a large collection of *URA3*-based cloning vectors of various types for yeast; the small size and known sequence of the *URA3* gene; the availability of numerous well-studied mutations in the *URA3* gene; and the specificity, case, and efficacy of the selection [18].

The acetate concentration of the fermentation medium had a great effect on pyruvate production by the mutants derived from chemical mutagenesis [15], but not by the disruptants derived from gene disruption (data not shown). These results showed that two species of mutants with reduced PDC activities have difference metabolic flux consequences. The former needed to add acetate to synthesize the adequate acetyl coenzyme A (acetyl-CoA) by acetyl-CoA synthetase (ACS) to reduce the degradation of pyruvate. Thus, more pyruvate could be accumulated after adding acetate during fermentation. The disruptants have no such behavior; consequently there will be a simpler pyruvate production process.

The essential role of PDC in assimilatory carbon metabolism is not a general phenomenon among yeasts. For example, it is an essential enzyme for the growth of *S. cerevisiae* on glucose, even in complex media; mutants of *Kluyveromyces lactis* lacking the unique *PDC1* strain lack pyruvate decarboxylase, but exhibit wild-type specific growth rates on glucose [37]. The role of PDC in *T. glabrata* was apparently different from that of the two yeasts *S. cerevisiae* and *K. lactis* (Fig. 4). However, the detailed mechanism needs to be explored further.

Pyruvate, the end-product of glycolysis, is located at the important branch-point between respiratory and fermentative carbon metabolism. Flux distribution at the pyruvate level is of crucial importance for by-product formation in yeast [32]. During fermentative sugar metabolism, pyruvate is decarboxylated to acetaldehyde by PDC, and the acetaldehyde is subsequently turned into ethanol by alcohol hydrogenase. Dissimilation of pyruvate for respiration requires its conver-

sion to acetyl-CoA, the precursor metabolite of the tricarboxylic acid (TCA) cycle, or to other metabolites of the TCA cycle. It is known that this conversion can occur three ways in *S. cerevisiae*: the direct oxidative pathway of pyruvate to acetyl-CoA via the mitochondrial PDH complex; the indirect oxidative pathway via PDC, acetaldehyde dehydrogenase (ADH) and ACS; and the anaplerotic carboxylation of pyruvate to oxaloacetate catalyzed by PYC [32]. In previous studies [6,15], the degradation of pyruvate was controlled either by a limited addition of thiamine, which is a cofactor of the PDH complex and the PDC enzyme in *T. glabrata*, or by screening mutants with reduced PDC-activity after random mutagenesis. Thiamine-limitation was efficient but not perfect for pyruvate fermentation by *T. glabrata*. Generally, the intracellular pyruvate concentration could be high, but it is soon channeled into ethanol production with increased rates of glycolysis in yeast [38]. In this study, *pdc* mutants derived from *T. glabrata* IFO005 exhibited the ability to produce increased levels of pyruvate than the parent strain (Table 3). This was because the *pdc* mutant had decreased PDC activity; the PDC activities of *pdc* mutants were only about 33% that of the parent strain. It was shown that the reduced PDC activity levels of *pdc* mutants would metabolize less pyruvate to ethanol via acetaldehyde. However, the levels of PDC activity in the parent strains (strains IFO005, IFO005-3, and IFO005-7) were too high to accumulate any significant level of pyruvate, the key intermediate metabolite in both respiration and fermentation. On the other hand, it was presumed that the level of the acetaldehyde-producing enzyme (PDC) and the acetaldehyde-degrading enzymes (ADH and A/DH) were well balanced in *pdc* mutants, which had reduced PDC enzyme levels. However, in the parent strain, the balance between these enzymes was not favorable for producing significant levels of pyruvate because its PDC level was too high. These studies demonstrated that it was possible to accumulate high levels of an intermediate metabolite such as pyruvate by altering the level of the PDC enzyme. According to the reported results, an efficient redirecting of metabolic flux to increase pyruvate production requires deletion of by-product synthesis pathways (i.e., ethanol, glycerol, etc.).

Improvement of strains has traditionally relied on random mutagenesis followed by screening for mutants with desirable properties. Now, the opportunity to disrupt related genes and regulatory elements distinguishes metabolic engineering from the traditional genetic approaches. In this work, the use of a straightforward metabolic engineering strategy consisting of *PDC* gene disruption in order to control the excessive conversion of intracellular pyruvate to ethanol increased pyruvate production in the pyruvate-producing yeast, *T. glabrata*. We can attribute the success of this strategy to previous detailed studies of carbon and electron flux [32] which revealed the catabolic behavior of *T. glabrata* with glucose substrates. Our efforts on the metabolic engineering of *T. glabrata* have achieved significant progress in improving the efficiency of the engineered strain in pyruvate production.

Acknowledgments

This work was supported by the National Knowledge Innovation Project from the Chinese Academy of Sciences. We thank Dr. Liu Jiquan, and Dr. H. Yde Steensma for their gifts of strain (*Torulopsis glabrata* IFO005) and plasmid pRUL129, respectively. We thank Dr. Keqian Yang (Institute of Microbiology, CAS) and Dr. Xiaorong Wu (Kansas State University, USA) for revising the manuscript.

References

[1] Yonehara T, Miyata R, Matsuno H, Goto M, Yahanda S. Development of fermentative production of pyruvate by metabolic control (Japanese). Seibutsu-Kogaku Kaishi 2000;78:56–62.
[2] Enei H, Nakazawa H, Matsui H, Okumura S, Yamada H. Enzymic preparation of L-tyrosine or 3, 4-dihydroxyphenyl-L-alanine from pyruvate, ammonia and indole. FEBS Lett 1972;21:39–41.
[3] Nakazawa H, Enei H, Okumura S, Yamada H. Synthesis of L-tryptophan from pyruvate, ammonia and indole. Agric Biol Chem 1972;36:2523–8.
[4] Kalman D, Colker CM, Wilets I, Roufs JB, Antonio J. The effects of pyruvate supplementation on body composition in overweight individuals. Nutrition 1999;15:337–40.
[5] Yokota A, Shimizu H, Terasawa Y, Takaoka N, Tomita F. Pyruvic acid production by lipoic acid auxotroph of *Escherichia coli* W1485. Appl Microbiol Biotechnol 1994;41:638–43.
[6] Yonehara T, Miyata R. Fermentative production of pyruvate from glucose by *Torulopsis glabrata*. J Ferment Bioeng 1994;78:155–9.
[7] Takao S, Tamids M. Pyruvic acid production by *Schizophyllum commune*. J Ferment Technol 1982;60:277–80.
[8] Zelić B, Gerharz T, Bott M, Vasić-Rački Đ, Wandrey C, Takors R. Fed-batch process for pyruvate production by recombinant *Escherichia coli* YYC 202 strain. Eng Life Sci 2003;3:299–305.
[9] Tomar A, Eiteman MA, Altman E. The effect of acetate pathway mutations on the production of pyruvate in *Escherichia coli*. Appl Microbiol Biotechnol 2003;62:76–82.
[10] Causey TB, Shanmugam KT, Yomano LP, Ingram LO. Engineering *Escherichia coli* efficient conversion of glucose to pyruvate. PNAS 2004;101:2235–40.
[11] Li Y, Hugenholtz J, Chen J, Lim SY. Enhancement of pyruvate production by *Torulopsis glabrata* using a two-stage oxgen supply control strategy. Appl Microbiol Biotechnol 2002;60:101–7.
[12] Zelić B, Gostović S, Vuorilehto K, Vasić-Rački Đ, Takors R. Process strategies to enhance pyruvate production with recombinant Escherichia coli: from repetitive fed-batch to in situ product recovery with fully integrated electrodialysis. Biotechnol Bioeng 2004;85:638–46.
[13] Hua Q, Shimizu K. Effect of dissolved oxygen concentration on the intracellular flux distribution for pyruvate fermentation. J Biotechnol 1999;68:135–47.
[14] Li Y, Chen J, Lun SY, Rui XS. Efficient pyruvate production by a multi-vitamin auxotroph of *Torulopsis glabrata*: key role and optimization of vitamin levels. Appl Microbiol Biotechnol 2001;55:680–5.
[15] Miyata R, Yonehara T. Breeding of high-pyruvate-producing *Torulopsis glabrata* with acquired reduced pyruvate decarboxlase. J Biosci Bioeng 1999;88:173–7.
[16] Bailey JE. Toward a science of metabolic engineering. Science 1991;252:1668–75.
[17] Sambrook JE, Fritsch EF, Maniatis T. Molecular cloning: a laboratory manual. 2nd ed. New York: Cold Spring Harbor; 1989. Cold Spring Harbor Laboratory.

[18] Boeke JD, Trueheart J, Natsoulis G, Fink GR. 5-Fluoroorotic acid as a selective agent in yeast molecular genetics. Methods Enzymol 1987;154:164–75.
[19] Burke D, Dawson D, Stearns T. Methods in yeast genetics. New York: Cold Spring Harbor; 2000. Cold Spring Harbor Laboratory.
[20] Berg MV, Steensma HY. Expression cassettes for formaldehyde and fluorocostate resistance, two dominant markers in *Saccharomyces cerevisiae*. Yeast 1997;13:551–9.
[21] Stockland AE, Clemente CL. Multiple forms of lactate dehydrogenase in *Staphlococcus aureus*. J Bacteriol 1969;100:347–53.
[22] Stellmach B. Bestimmungsmethoden enzyme für pharmazie, lebensmittelchemie, technil, biochemie, medizin. Steinkopff: Darmstadt:1987;pp22–3.
[23] Song HB, Demain AL. An improved colorimetric assay for polyols. Anal Biochem 1977;81:18–20.
[24] Ullrich J. Yeast pyruvic decarboxylase (2-oxoacid carboxyase, EC 4. 1. 1. 1) assay of thiamine pyrophosephate. Method Enzymol 1970;18:109–15.
[25] Zehender H, Trescher D, Ullrich J. Activity stain for pyruvate decarboxylase in polyacrylamide gels. Anal Biochem 1983;135: 16–21.
[26] Candy JM, Duggleby RG, Mattick JS. Expression of active yeast pyruvate decarboxylase in *Escherichia coli*. J Gen Microbiol 1991;137:2811–5.
[27] Bradford MM. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 1976;72:248–54.
[28] Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 1970;227:680–5.
[29] Appleby T, Kinsland C, Begley TP, Ealick SE. The crystal structure and mechanism of orotidine 5′-monophosphate decarboxylase. PNAS 2000;97:2005–10.
[30] Wu N, Mo Y, Gao J, Pai EF. Electrostatic stress in catalysis: structure and mechanism of the enzyme orotidine monophosphate decarboxylase. PNAS 2000;97:2017–22.
[31] Gonzalez C, Perdomo G, Tejera P, Brito N, Siverio JM. One-step, PCR-mediated, gene disruption in the yeast *Hanzenula polymorpha*. Yeast 1999;15:1323–9.
[32] Pronk J, Yde Steensma H, Van Dijken JP. Pyruvate metabolism in *Saccharomyces cerevisiae*. Yeast 1996;12:1607–33.
[33] Wang Q, He P, Lu D, Shen A, Jiang N. Screening of pyruvate-producing yeast and effect of nutritional conditions on pyruvate production. Lett Appl Microbiol 2002;35:338–42.
[34] Miyata R. Production of optically active alpha-hydroxyketone. 2000; JP2000093189.
[35] Zhou P, Szczypka MS, Young R, Thiele DJ. A system for gene cloning and manipulation in the yeast *Candida glabrata*. Gene 1994;142:135–40.
[36] Mehra RK, Thorvaldsen JL, Macreadie IG, Winge DR. Cloning system for *Candida glabrata* using elements from the metallothionein-encoding gene that confer autonomous replication. Gene 1992;113:119–24.
[37] Flikweert MT, de Swaaf M, van Dijken J, Pronk JT. Growth requirements of pyruvate-decarboxylase-negative *Saccharomyces cerevisiae*. FEMS Microbiol Lett 1999;174:73–9.
[38] Piolakis ES, Bartley W. Change in the enzyme activities of *Saccharomyces cerevisiae* during aerobic growth of yeast on different carbon sources. Biochem J 1965;97:284–97.

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Metabolic Engineering 7 (2005) 27–37
METABOLIC ENGINEERING
www.elsevier.com/locate/ymben

Semi-synthetic DNA shuffling of *aveC* leads to improved industrial scale production of doramectin by *Streptomyces avermitilis*

Kim Stutzman-Engwall[a,*], Steve Conlon[a], Ronald Fedechko[a], Hamish McArthur[a], Katja Pekrun[b], Yan Chen[b], Stephane Jenne[c], Charlene La[c], Na Trinh[c], Seran Kim[c], Ying-Xin Zhang[c], Richard Fox[c], Claes Gustafsson[b], Anke Krebber[c]

[a]*Pfizer Global Research & Development, Gorton Laboratories, Eastern Pt. Rd. MS 4123, Groton, CT 06340, USA*
[b]*Maxygen, 515 Galveston Drive, Redwood City, CA 94063, USA*
[c]*Codexis, 200 Penobscot Drive, Redwood City, CA 94063, USA*

Received 6 April 2004; accepted 7 July 2004
Available online 22 January 2005

Abstract

The avermectin analog doramectin (CHC-B1), sold commercially as Dectomax™, is biosynthesized by *Streptomyces avermitilis*. *aveC*, a gene encoding an unknown mechanistic function, plays an essential role in the production of doramectin (avermectin CHC-B1), modulating the production ratio of CHC-B1 to other avermectins, most notably the undesirable analog CHC-B2. To improve the production ratio for doramectin, the *aveC* gene was subjected to iterative rounds of semi-synthetic DNA shuffling. Libraries of shuffled *aveC* gene variants were transformed into *S. avermitilis*, screened using a miniaturized 96-well growth and production format, and analyzed by high throughput mass spectrometry to determine CHC-B2:CHC-B1 ratios. Several improved *aveC* variants were identified; the best shuffled gene encoded 10 amino acid mutations, and conferred a final CHC-B2:CHC-B1 ratio of 0.07:1, a 23-fold improvement over the starting gene (*aveC* wild type). Chromosomal insertion of an improved *aveC* shuffled gene into a high titer *S. avermitilis* strain yielded an improved doramectin production strain. This strain is under development to be used commercially, and is expected to provide considerable cost savings in large-scale manufacture, as well as significantly reducing by-product levels of CHC-B2 requiring disposal.

Keywords: *Streptomyces avermitilis*; *aveC*; Semi-synthetic DNA shuffling

1. Introduction

Polyketides are a large and structurally diverse group of natural products with activities ranging from antibacterial (erythromycin), anticancer (daunorubicin), antifungal (amphotericin), cholesterol-lowering (mevastatin), immunosuppressant (rapamycin) to anthelmintic (avermectin). The highly complex polyketide carbon chain backbone is produced by the sequential activity of the multi-enzyme and often multi-modular polyketide synthase. Variations in the number of extensions, choice of starter unit and extender units, stereochemistry, reduction of the polyketide chain, and subsequent modification of the polyketide backbone by methyl-, sugar-, and hydroxyl groups, all contribute to the structural diversity of these types of compounds (Katz and Donadio, 1993). As a class, polyketides are one of the richest sources of pharmaceuticals today. Consequently, considerable current research is focused on manipulation and design of polyketide synthetic pathways (Chartrain et al., 2000).

Avermectins are 16-membered macrocyclic lactones produced by *Streptomyces avermitilis* (Burg et al., 1979). The natural and semi-synthetic derivates of avermectin

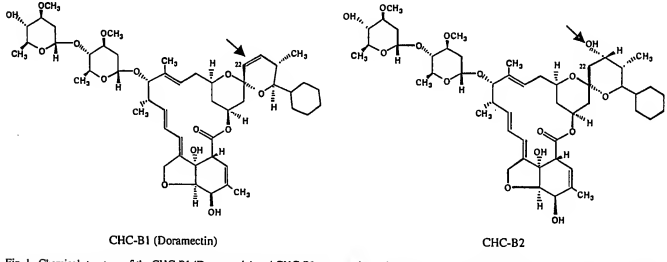
*Corresponding author.
E-mail address: stutzmanengw@groton.pfizer.com (K. Stutzman-Engwall).

1096-7176/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymben.2004.07.003

are potent anthelmintic and insecticidal compounds and are widely used as veterinary treatments for broad spectrum parasite control and in human medicine for the treatment of river blindness (onchocerciasis). Avermectin derivatives are the most widely used drugs in animal health and agriculture, with current worldwide sales exceeding 1 billion US dollars. The avermectin polyketide chain is derived from seven acetates and five propionates, together with a single 2-methylbutyric acid or isobutyric acid residue which forms the *sec*-butyl or isopropyl group attached to C25 (Cane et al., 1983; Chen et al., 1989). The polyketide chain is subsequently cyclized, glycosylated at C13 with the attachment of two oleandrose sugars and methylated at the hydroxyl group of C5. In the biosynthesis of the analogue doramectin, cyclohexanecarboxylic acid (CHC) is used in place of the natural chain initiator. The doramectin production strain is a mutant of *S. avermitilis* lacking branched-chain-2-oxo acid dehydrogenase (*bkd*) and *O*-methyltransferase activity. This strain produces only the B forms of avermectin corresponding to the supplemented fatty acid (Dutton et al., 1991; Hafner et al., 1991).

The *S. avermitilis* doramectin production strain synthesizes a mixture of two related compounds, CHC-B1 (doramectin) and CHC-B2. The CHC-B1 compound has a double bond between C22 and C23 whereas CHC-B2 contains a saturated C–C bond with a hydroxyl group at C23 (Fig. 1), and it has been proposed that the CHC-B2 form is converted to the B1 form by a dehydration event at C22,23 (Chen and Inamine, 1989). CHC-B2 is a less effective pharmaceutical agent than CHC-B1 (Goudie et al., 1993). Although characterization of the genes in the avermectin gene cluster has clarified much of the biosynthetic pathway (Ikeda et al., 1999; Ikeda and Omura, 1995; MacNeil et al., 1992), the mechanism that determines the production ratio CHC-B2:CHC-B1 remains unclear. The presumptive dehydration branch point resulting in the differentiation between the B1- and B2-type compounds is proposed to be an early event in the biosynthetic pathway; probably occurring during the PKS catalyzed condensation (Chen and Inamine, 1989). We have determined that the dehydratase function in module 2 of the PKS that would correspond to C22–23 dehydration does not have a role in determining CHC-B2:CHC-B1 ratio (Stutzman-Engwall et al., 1997). In contrast, when a portion of the *aveC* gene was randomly mutated, a single avermectin component B2a was produced (Ikeda et al., 1995). Deletion of *aveC* abolishes production of significantly detectable amounts of both avermectins B1 and B2, whereas overexpression of *aveC* has no effect on ratio in wild-type strains (Stutzman-Engwall et al., 2003). This *aveC* deletion mutant can be fully complemented *in trans*. Several unrelated mutations were identified after site-specific mutagenesis and error-prone PCR that resulted in significantly improved ratios of doramectin to CHC-B2 produced (Stutzman-Engwall et al., 2003). The majority of the observed random mutations, however, influence the production ratio adversely (in favor of CHC-B2).

We report here further improvements in production ratios obtained by evolving *aveC* by semi-synthetic DNA shuffling and screening for variants producing increasing amounts of doramectin relative to CHC-B2. This semi-synthetic shuffling format improved the distribution of recombinant genes in shuffled libraries, which reduced the number of variants that needed to be screened to identify improved variants. The best variants identified from the first round of semi-synthetic shuffling produced a CHC-B2:CHC-B1 ratio of 0.2:1, while retaining the same total avermectin (CHC-B2 + CHC-B1) yield as the wild-type gene. Subsequent rounds of



CHC-B1 (Doramectin) CHC-B2

Fig. 1. Chemical structure of the CHC-B1 (Doramectin) and CHC-B2 avermectin analogues. Arrows mark the enoyl carbons at C22-23 in the CHC-B1 analogue and the hydroxyl group at C23 in the CHC-B2 analogue.

semi-synthetic gene shuffling yielded further improvements, resulting in an *aveC* variant with a CHC-B2:CHC-B1 ratio of 0.07:1. The best *aveC* gene was subsequently introduced into the chromosome of the production strain; this new production strain achieved a maximal ratio improvement.

2. Materials and methods

2.1. General methods

Escherichia coli DH5α was used for routine subcloning employing standard growth media and conditions (Sambrook et al., 1989), or as described in suppliers' protocols. *E. coli* transformants were selected with 100 μg/ml ampicillin. *S. avermitilis aveC* mutant SE180-11 (Stutzman-Engwall et al., 2003) containing an insertionally inactivated *aveC* gene, was used as host for complementation experiments. Plasmid pWHM3 (Vara et al., 1989) was used for routine subcloning for transformation into *S. avermitilis*. Plasmids were isolated from *S. lividans* or *E. coli* DM1 (BRL) or *E. coli* SCS110 (Stratagene) prior to transformation into *S. avermitilis*. *S. avermitilis* routine and high throughput (HTP) culture conditions were similar to those described previously (Stutzman-Engwall et al., 2003). *S. avermitilis* was grown for fermentation, and assays performed as previously described (Pacey et al., 2000; Stutzman-Engwall et al., 2001). DNA was sequenced using a *Taq* Dye Deoxy terminator cycle sequencing kit and an ABI 373A Automated DNA Sequencing system (Perkin Elmer, Foster City, CA). Sequence data were assembled and edited using Genetic Computer Group programs (GCG, Madison, WI, Devereux et al., 1984). PCR amplification and genomic replacement protocols were similar to those described previously (Stutzman-Engwall et al., 2003).

2.2. Plasmid and strain constructions

S. avermitilis aveC mutant SE247-11 (Stutzman-Engwall et al., 2003) was used as the production strain host to introduce gene replacement vectors containing mutated *aveC* DNA. Gene replacement vector pSE370 was constructed in a 4-way ligation containing a 0.6 kb *SphI*/*PstI* fragment from pSE617 (CHC-B2:CHC-B1 ratio of 0.05:1), a 1.4 kb *PstI*/*Bam*HI fragment from pSE180, a 2.8 kb *XbaI*/*Sph*I fragment from pSE180 and *Bam*HI/*Xba*I-cleaved pWHM3. Because pSE370 was missing *aveC* mutation E238D, a second gene replacement vector pSE375 was constructed. pSE370 was used as a template to PCR a ~600 bp *SphI*/*NsiI* fragment using a mismatched primer (5′-CTGCTGCATGCTGGGCTCGCTGCGCTTCTTCCGCGACGACC-3′) and reverse primer (5′-GCTGGGAAACCAAGGGATCG-3′). The PCR product was digested with *SphI*/*NsiI* and a 3-way ligation with a 9.2 kb *SphI*/*Hind*III fragment and a 2.5 kb *Hind*III/*NsiI* fragment was performed. All plasmid constructions were confirmed by restriction analysis and, in some cases, DNA sequence analysis.

2.3. Semi-synthetic shuffling

Previously, mutation of the *aveC* gene yielded 7 variants, each conferring a reduced CHC-B2:CHC-B1 ratio while maintaining overall avermectin yield (Stutzman-Engwall et al., 2003). DNA sequence analysis determined that each clone encoded between two and four AA substitutions. The best first round clone (pSE290, CHC-B2:CHC-B1 ratio of 0.4:1) was shuffled using a method described by Stemmer (1994b). Individual oligonucleotides encoding AA substitutions corresponding to single mutations conferring an improved B2:B1 product ratio were added to the shuffling reaction at 5 molar excess over *aveC* template DNA. In the case of AA136, AA138, and AA139, pairs of AA substitutions were constructed in a single oligonucleotide. Each nucleotide mismatch in an oligonucleotide was flanked by 20 nucleotides of perfect identity to optimize incorporation during the shuffling reaction. Oligonucleotides were purchased from Operon technologies (Alameda, CA). Subsequent libraries were generated using best available full length *aveC* variants as the shuffling template, and oligonucleotides encoding other beneficial mutations were added to the reaction. Mutazyme (Stratagene) was added to the reaction when additional diversity from random DNA mutations was desired.

2.4. Growth of clones for HTP analysis of doramectin production

Shuffled libraries were cloned into pWHM3 and transformed into methylation negative *E. coli* INV110 cells (Invitrogen). DNA sequencing of at least 10 randomly chosen colonies verified the quality of a shuffled library. Plasmid DNA from ~10^5 pooled clones was isolated and transformed into *S. avermitilis* SE180-11 protoplasts. Individual clones were grown at 30 °C (200 μL R5 medium (Kieser et al., 2000) containing 5 μg/ml thiostrepton in deep 96-well seed plates in a humidified shaker at 250 rpm). A glass bead was included in each well for dispersion of mycelia and agitation of the culture. After 3–5 days, 20 μl of the seed culture was transferred to production plates. The remainder of the seed plate was stored as a master plate at 4 °C. Production plate media were prepared essentially as described in Pacey et al. (2000), with the addition of 1% agarose. The 96 well production plate (containing 1 ml production medium per well) was tilted during solidification of the medium to increase surface

area for growth. Inoculated production plates were incubated at 30 °C under humidity for 12–14 days.

2.5. Extraction and ESI-MS/MS screening

Ethyl acetate (1 ml) was added to each well, and the plate was shaken at room temperature for 20 min. Approximately 750 µl of the ethyl acetate phase was recovered, transferred to a 96-well plate, and allowed to evaporate over night. The residue was resuspended in 100 µl methanol containing 1 mM NaCl; 5 µl was injected directly into a triple-quadrupole mass spectrometer (Micromass Quattro LC) from the 96-well format using a TwinPal LEAP auto sampler. Products were analyzed in flow injection mode without prior liquid chromatography separation. This method allowed rapid (<25 s per sample), quantitative and precise detection of products by detecting the molecular weight of the product, as well as a characteristic fragmentation ion produced in the third quadrupole of the mass spectrometer. The MS/MS transition for CHC-B1 sodiated ion is from m/z 921 to m/z 777; for CHC-B2, the sodiated ion is from m/z 939 to m/z 795 in positive mode. Integration of the separate CHC-B1 and CHC-B2 chromatograms for each well identified clones with altered CHB-B2 over CHB-B1 ratios. Clones producing improved ratios were retested in triplicate starting from the original seed plate using the same growth, production and assay protocol as described above. Clones producing improved ratios were confirmed by fermentation in 30 ml shake flask, under conditions described previously (Stutzman-Engwall et al., 2001).

3. Results

3.1. Semi-synthetic shuffling of improved avec clones

In certain previously described formats for DNA shuffling, each improved gene is isolated and fragmented, and a mixture of fragments from each gene is recombined to create the next round library (Stemmer, 1994a, b). Recombination of variants by DNA shuffling creates novel permutations of substitutions, however, to adequately explore the potential genetic diversity of a given library, a very large number of variants from that library must be screened (Moore et al., 1997). In particular, to identify clones possessing higher numbers of substitutions conferred by distinct parental gene sequences, large libraries may either be screened extensively, or multiple iterative rounds or DNA shuffling and screening of smaller numbers of variants may be pursued.

In the present study, due in part to long growth periods required for *S. avermitilis*, the screen is both complex and time-consuming (a single round of shuffling, transformation, growth, production and screening of 5000 variants, followed by characterization of potential hits can require approximately 2–3 months). The implementation of a new, semi-synthetic shuffling (oligonucleotide-based) format (Crameri et al., 2000) allowed for the exploration of diversity afforded by all variations identified in the initial screen of first round variants in a much smaller library. The theory behind the semi-synthetic approach pursued here is similar to other techniques that can be used to accelerate the recombination of beneficial diversity (Aita et al., 2002; Vajdos et al., 2002).

Ten amino acid substitutions present in the seven first round hits generated by error prone mutagenesis corresponded with improved CHC-B2:CHC-B1 ratios (Stutzman-Engwall et al., 2003). Oligonucleotides encoding individual substitutions (A61T, F99S, G111V, L136P, G179S, V196A, E238D, P289L; plus all combinations of L136P/S138T/A139T and A139F) were designed so that all combinations of amino acid changes would be afforded with equal probability in the resulting library. Oligonucleotides were mixed and shuffled with fragmented *aveC* derived from the best first round variant, pSE290 (D48E, A89T), which produced a CHC-B2:CHC-B1 ratio of 0.4:1.

Fifteen randomly isolated clones from the library were sequenced to determine oligonucleotide incorporation and error frequency. All but one of the substitutions were incorporated in 12 clones at least once (the average number of mutations incorporated was 4.5). In addition, 3 random (non-oligonucleotide-derived) point mutations encoding an amino acid change were identified in 3636 codons (mutation rate of 0.08%). Approximately 5000 clones derived from the semi-synthetic shuffled library were screened for production of CHC-B1 and CHC-B2. Approximately 10% of colonies inoculated on the production plate did not grow, and approximately 20% of the analyzed clones did not produce CHC-B1 yields above that of the *aveC* deletion mutant. These percentages are consistent with those previously reported for this assay (Stutzman-Engwall et al., 2003). Among the remaining 70%, 89 were identified as producing a CHC-B2:CHC-B1 ratio that was improved over the previously identified best clone (Fig. 2). A 70% live rate of active variants in the library corresponds to a 3.5-fold oversampling of the theoretical library size ($2^{10} = 1024$), resulting in approximately 97% coverage of unique variants (Patrick et al., 2003).

The 89 promising variants were evaluated in production tests (30 ml liquid media). Duplicate HPLC testing determined that 90% of these variants produced a CHC-B2:CHC-B1 ratio in liquid growth that was consistent with the first tier assay (96 well plate using solid media), validating the miniaturized assay format and demonstrating the correlation between solid and liquid growth screen, and a correlation between detection by MS and

HPLC. Among these, eight variants reproducibly produced a CHC-B2:CHC-B1 at a ratio of 0.2:1.

DNA sequencing revealed that each of the eight improved second round variants were unique, encoding on average seven amino acid changes (Fig. 3, Table 1). The success of DNA shuffling relies on rapid fixation and accumulation of cumulative beneficial mutations from multiple parents, as well as the removal of deleterious mutations from the population (F99S, G111V). While the standard shuffling format is strongly biased toward the incorporation of fewer, potentially beneficial mutations, with the most likely number of mutations at 1 per gene, the semi-synthetic format demonstrates a more uniform probability distribution, with the most likely number of mutations at 5 per gene. In particular, the probability of incorporating 7 mutations per gene as seen here is over 10,000 times more likely when using the semi-synthetic approach, as compared to standard DNA shuffling. Furthermore, semi-synthetic shuffling will allow for uncoupling of mutations that are positioned close in sequence space and are unlikely to be recombined by traditional crossover events. In practice, a random mutation rate on the order of ~0.1% per base pair was observed in addition to changes introduced by oligonucleotides (this includes silent codon changes having no influence on the amino acid level).

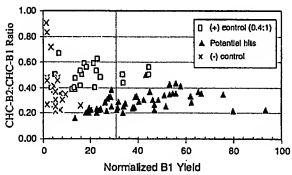


Fig. 2. Data points taken from the assay of a 96-well plate containing improved ratio variants after semi-synthetic DNA shuffling. The result is displayed as the ratio of CHC-B2:CHC-B1 as a function of the yield of CHC-B1. Clones were identified that produced CHC-B2:CHC-B1 ratios better than 0.4:1. The cut-off value for active clones was determined based on the highest yield CHC-B1 achieved from the negative control *aveC* deletion strain averaged over several experiments. Only clones producing >30 μg/ml normalized B1 (vertical line) were analyzed further (normalized data derived from μg/ml production levels). The positive control is *S. avermitilis* strain SE180-11 transformed with pSE290 (D48E, A89T). The negative control is *S. avermitilis* strain transformed with the vector pWHM3.

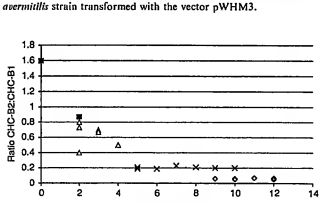


Fig. 3. CHC-B2:CHC-B1 ratio as a function of the number of substitutions per clone as compared to the wildtype *aveC* gene. Squares represent the two starting points; wildtype *aveC* and the S138T, A139T variant, pSE231. Open triangles represent the 7 best hits from the first round, X represent the 8 best hits from the second round, and open diamonds represent the 9 best hits from the third round.

3.2. Generation of additional genetic diversity

Semi-synthetic gene shuffling yielded a significant improvement in the CHC-B2:CHC-B1 ratio over clone pSE290 (0.4:1 to 0.2:1). Sequencing of 75 independent *aveC* clones from the semi-synthetic DNA shuffling reaction determined that at least two mutations were conserved in all cases (D48E and G179S, Table 1). It was not likely that the mutations introduced by semi-synthetic shuffling represented all possible beneficial genetic diversity that could be obtained in the *aveC* gene, so additional diversity was realized by constructing and screening additional libraries of randomly mutated *aveC* clones. This was accomplished by repeating error prone amplification of wt *aveC* in the presence of 0.2 mM manganese (Stutzman-Engwall et al., 2003) or Mutazyme. Approximately 3000 clones derived from these libraries were screened for production of CHC-B1 and CHC-B2; of these, 50 variants produced a CHC-B2:CHC-B1 ratio better than 1.3:1.

These 50 gene variants were evaluated in production tests (30 ml liquid media) and tested in duplicate by HPLC. Of these variants, 42 produced detectable CHC-B1 and 26 produced a CHC-B2:CHC-B1 ratio in liquid growth that was at least 1.3:1 or better (Fig. 4). DNA sequence analysis of the best 23 *aveC* clones determined that significant new genetic diversity was present. Several clones possessed a single mutation in *aveC* that resulted in an improved CHC-B2:CHC-B1 ratio of ~1:1. Ten new amino acid substitutions were identified that conferred improved CHC-B2:CHC-B1 ratios (S41G, R71L, L87V, W110L, L136M, T149S, F176C, L206M, G209R, and I280V). Oligonucleotides encoding these substitutions were designed so that all combinations of amino acid changes would be provided with equal probability. Oligonucleotides were combined and shuffled together with fragmented *aveC* derived from the best second round variant, pSE554 (Table 1).

3.3. Second and third rounds of semi-synthetic gene shuffling

Approximately 5000 clones from the second round of semi-synthetic shuffling were screened for production of

Table 1
Amino acid sequence of the best eight *aveC* variants generated after the first round of semi-synthetic DNA shuffling. The parent *aveC* gene used in the shuffling reaction was pSE290 and contained the mutations D48E and A89T. CHC-B2:CHC-B1 ratios are pSE506 0.19:1; pSE549 0.17:1; pSE554 0.18:1, pSE537 0.19:1; pSE505 0.2:1; pSE502 0.2:1; pSE503 0.19:1; pSE515 0.19:1

```
      1                                                                100
pSE506  VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQQGGSKDTGTTEVVYPVISVVCITAAAAWLFRRCRVERRLLFDALLFLGLLFTSWQSPLMNWFH
pSE549  VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQQGGSKDTGTTEVVYPVISVVCITAAAAWLFRRCRVERRLLFDALLFLGLLFTSWQSPLMNWFH
pSE554  VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQQGGSKDTGTTEVVYPVISVVCITTAAAWLFRRCRVERRLLFDALLFLGLLFTSWQSPLMNWFH
pSE537  VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQQGGSKDTGTTEVVYPVISVVCITTAAAWLFRRCRVERRLLFDALLFLGLLFTSWQSPLMNWFH
pSE505  VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQQGGSKDTGTTEVVYPVISVVCITAAAAWLFRRCRVERRLLFDALLFLGLLFTSWQSPLMNWFH
pSE502  VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQQGGSKDTGTTEVVYPVISVVCITAAAAWLFRRCRVERRLLFDALLFLGLLFTSWQSPLMNWFH
pSE503  VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQQGGSKDTGTTEVVYPVISVVCITAAAAWLFRRCRVERRLLFDALLFLGLLFTSWQSPLMNWFH
pSE515  VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQQGGSKDTGTTEVVYPVISVVCITAAAAWLFRRCRVERRLLFDALLFLGLLFTSWQSPLMNWFH
aveC WT VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQQGGSKDTGTTDVVYPVISVVCITAAAAWLFRRCRVERRLLFDALLFLGLLFASWQSPLMNWFH

      101                                                              200
pSE506  SVLVSNASVWGAVGSWGPYVPGWQAGPGAEAEMPLATESVCMSALIVTVLCSKALGWIKARRPAWRTWRLVLAVFFISIVLGLSEPLPSASGISAWARA
pSE549  SVLVSNASVWGAVGSWGPYVPGWQAGPGAEAEMPLATTSVCMSALIVTVLCSKALGWIKARRPAWRTWRLVLAVFFISIVLGLSEPLPSASGISAWARA
pSE554  SVLVSNASVWGAVGSWGPYVPGWQAGPGAEAEMPLATTSVCMSALIVTVLCSKALGWIKARRPAWRTWRLVLAVFFISIVLGLSEPLPSASGISAWARA
pSE537  SVLVSNASVWGAVGSWGPYVPGWQAGPGAEAEMPLATTSVCMSALIVTVLCSKALGWIKARRPAWRTWRLVLAVFFISIVLGLSEPLPSASGISAWARA
pSE505  SVLVSNASVWGAVGSWGPYVPGWQAGPGAEAEMPLATTSVCMSALIVTVLCSEALGWIKARRPAWRTWRLVLAVFFISIVLGLSEPLPSASGISAWARA
pSE502  SVLVSNASVWGAVGSWGPYVPGWQAGPGAEAEMPEASASVCMSALIVTVLCSKALGWIKARRPAWRTWRLVLAVFFISIVLGLSEPLPSASGISVWARA
pSE503  SVLVSNASVWGAVGSWGPYVPGWQAGPGAEAEMPEASASVCMSALIVTVLCSEALGWIKARRPAWRTWRLVLAVFFISIVLGLSEPLPSASGISVWARA
pSE515  SVLVSNASVWGAVGSWGPYVPGWQAGPGAEAEMPEASASVCMSALIVTVLCSKALGWIKARRPAWRTWRLVLAVFFISIVLGLSEPLPSASGISVWARA
aveC WT SVLVSNASVWGAVGSWGPYVPGWQAGPGAEAEMPLASASVCMSALIVTVLCSKALGWIKARRPAWRTWRLVLAVFFIGIVLGLSEPLPSASGISVWARA

      201                                                              303
pSE506  LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDRDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHILLSLVGGQPPDQLPDSFQAPAAY*
pSE549  LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDRDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHILLSLVGGQEPDQLPDSFQAPAAY*
pSE554  LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDRDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHILLSLVGGQLPDQLPDSFQAPAAY*
pSE537  LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHILLSLVGGQLPDQLPDSFQAPAAY*
pSE505  LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDRDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHILLSLVGGQLPDQLPDSFQAPAAY*
pSE502  LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDRDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHILLSLVGGPPDQLPDSFQAPAAY*
pSE503  LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDRDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHILLSLVGGPPDQLPDSFQAPAAY*
pSE515  LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDERDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHILLSLVGGQPPDQLPDSFQAPAAY*
aveC WT LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDERDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHILLSLVGGQPPDQLPDSFQAPAAY*
```

CHC-B1 and CHC-B2; 67 variants produced a CHC-B2:CHC-B1 ratio better than 0.2:1. These improved variants were progressed through shake flask validation in liquid medium and tested in duplicate, as described above; 88% of these variants produced a CHC-B2:CHC-B1 ratio consistent with the first tier solid media assay. Of these, six reproducibly produced CHC-B2:CHC-B1 at a ratio of 0.1:1 (Fig. 5, Table 2). DNA sequence analysis determined that the clones contained a unique complement of mutations, although all six clones had acquired the mutations D48E, A89T, T149S, G179S, and E238D.

Further analysis of the variants from the random mutation library determined that three new amino acid substitutions conferred improved CHC-B2:CHC-B1 ratios (V2M, S90N, C142Y). Oligonucleotides encoding these substitutions, as well as the oligonucleotides from the second round of semi-synthetic shuffling, were designed so that all combinations of amino acid changes could be allowed with equal probability. The oligonucleotides were combined and shuffled with fragmented *aveC* derived from pSE554. Approximately 5000 clones from the third round of semi-synthetic shuffling were screened for production of CHC-B1 and CHC-B2; 96 were identified as producing a CHC-B2:CHC-B1 ratio better than 0.1:1.

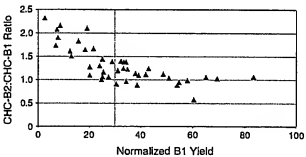


Fig. 4. Data points taken after fermentation analysis of *aveC* variants at 30 ml scale after a new round of random mutagenesis to identify new genetic diversity. The result is displayed as the ratio of CHC-B2:CHC-B1 as a function of the yield of CHC-B1 (normalized data derived from μg/ml production levels). The best 23 clones (CHC-B2:CHC-B1 ratios better than 1.3:1) were sequenced.

The 96 improved variants were retested in liquid media, as described before. Eighty percent of the variants produced a CHC-B2:CHC-B1 ratio in liquid medium that was consistent with the solid media assay results. DNA sequence analysis was completed on the best 16 variants, all of which reproducibly produced CHC-B2:CHC-B1 at a ratio of 0.07:1 or less (Fig. 5, Table 3). Although each sequenced clone contained a unique complement of mutations, all sixteen clones had acquired the common mutations D48E, R71L, A89T, T149S, F176C, G179S, and E238D.

3.4. Insertion of the CHC-B2:CHC-B1 0.07:1 mutation into the S. avermitilis production strain chromosome

Two different approaches for gene replacement vector construction were chosen. The first approach used the *aveC* internal restriction fragment *PstI*/*SphI* (~660 bp) from the mutated *aveC* clone pSE617 to subclone into the gene replacement vector, as previously described (Stutzman-Engwall et al., 2003). The simplicity of this approach is attractive; however, mutations outside of the *SphI*/*PstI* fragment are not retained (see Table 3). The gene replacement plasmid pSE370 (containing the *aveC* gene with mutations D48E, A61T, R71L, A89T, L136P, T149S, F176C, G179S, and V196A) was transformed into protoplasts of *S. avermitilis* SE247-11 (a high titer *aveC* knockout strain) and analyzed by fermentation for avermectin production. Fermentation analysis showed that the pSE370 transformants were producing very low avermectin titers. A second gene replacement strain was constructed to introduce the mutation E238D (which was 3′ to the *SphI* site and not included in the subcloning of pSE370). The gene replacement plasmid pSE375 (containing the *aveC* gene variant with mutations D48E, A61T, R71L, A89T, L136P, T149S, F176C, G179S, V196A, and E238D) was

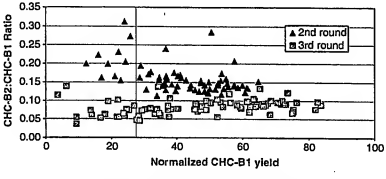


Fig. 5. Data points taken after fermentation analysis of improved *aveC* variants at 30 ml scale after the second and third rounds of semi-synthetic gene shuffling. Clones were identified that produced CHC-B2:CHC-B1 ratios better than 0.2:1 for second round (triangles) and better than 0.1:1 for the third round (squares). The cut-off value for active clones (vertical line) was determined based on the highest yield CHC-B1 achieved from the negative control *aveC* deletion strain averaged over several experiments (normalized data derived from μg/ml production levels).

Table 2

Amino acid sequence of the six best *aveC* variants generated after second round of semi-synthetic DNA shuffling. The parent *aveC* gene used in the shuffling reaction was pSE554 and contained the mutations D48E, A61T, S138T, A139T, G179S, V196A, E238D, and P289L. CHC-B2:CHC-B1 ratios are pSE603 0.13:1; pSE601 0.12:1; pSE607 0.13:1, pSE609 0.13:1; pSE608 0.12:1; pSE610 0.13:1

```
           1                                                                                                  100
pSE603     VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQGQGGSKDTGTTEVVYPVISVVCITAAAAWLFRRCLVERRLLFDALLFLGLLFTSWQSPLMNWFH
pSE601     VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQGQGGGKDTGTTEVVYPVISVVCITTAAAWLFRRCLVERRLLFDALLFLGLLFTSWQSPLMNWFH
pSE607     VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQGQGGSKDTGTTEVVYPVISVVCITTAAAWLFRRCLVERRLLFDALLFLGLLFTSWQSPLMNWFH
pSE609     VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQGRGGSKDTGTTEVVYPVISVVCITTAAAWLFRRCLVERRLLFDALLFLGLVFTSWQSPLMNWFH
pSE608     VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQGQGGSKDTGTTEVVYPVISVVCITTAAAWLFRRCLVERRLLFDALLFLGLLFTSWQSPLMNWFH
pSE610     VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQGQGGSKDTGTTEVVYPVISVVCITTAAAWLFRRCRVERRLLFDALLFLGLVFTSWQSPLMNWFH
aveC WT    VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQGQGGSKDTGTTDVVYPVISVVCITAAAAWLFRRCRVERRLLFDALLFLGLLFASWQSPLMNWFH

           101                                                                                                200
pSE603     SVLVSNASVWGAVGSWGPYVPGWQGAGPGAEAEMPPASASVCMSALIVSVLCSKALGWIKARRPAWRTWRLVLAVCFISIVLGLSEPLPSASGISVWARA
pSE601     SVLVSNASVWGAVGSWGPYVPGWQGAGPGAEAEMPMATTSVCMSALIVSVLCSKALGWIKARRPAWRTWRLVLAVFFISIVLGLSEPLPSASGISAWARA
pSE607     SVLVSNASVWGAVGSWGPYVPGWQGAGPGAEAEMPMATTSVCMSALIVSVLCSKALGWIKARRPAWRTWRLVLAVFFISIVLGLSEPLPSASGISAWARA
pSE609     SVLVSNASVWGAVGSWGPYVPGWQGAGPGAEAEMPMATTSVCMSALIVSVLCSKALGWIKARRPAWRTWRLVLAVFFISIVLGLSEPLPSASGISAWARA
pSE608     SVLVSNASVWGAVGSWGPYVPGWQGAGPGAEAEMPLASTSVCMSALIVSVLCSKALGWIKARRPAWRTWRLVLAVCFISIVLGLSEPLPSASGISAWARA
pSE610     SVLVSNASVLGAVGSWGPYVPGWQGAGPGAEAEMPLATTSVCMSALIVSVLCSKALGWIKARRPAWRTWRLVLAVFFISIVLGLSEPLPSASGISAWARA
aveC WT    SVLVSNASVWGAVGSWGPYVPGWQGAGPGAEAEMPLASASVCMSALIVTVLCSKALGWIKARRPAWRTWRLVLAVFFIGIVLGLSEPLPSASGISVWARA

           201                                                                                                    303
pSE603     LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDDRDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHVLLSLVGGQPPDQLPDSFQAPAAY*
pSE601     LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDDRDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCLHILLSLVGGQLPDQLPDSFQAPAAY*
pSE607     LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDDRDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHILLSLVGGQLPDQLPDSFQAPAAY*
pSE609     LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDDRDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHILLSLVGGQLPDQLPDSFQAPAAY*
pSE608     LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDDRDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHILLSLGGGQLPDQLPDSFQAPAAY*
pSE610     LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDDRDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHILLSLVGGQLPDQLPDSFQAPAAY*
aveC WT    LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDERDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHILLSLVGGQPPDQLPDSFQAPAAY*
```

Table 3
Amino acid sequence of the nine best *aveC* variants generated after the third round of semi-synthetic DNA shuffling. The parent *aveC* gene used in the shuffling reaction was pSE554 and contained the mutations D48E, A61T, S138T, A139T, G179S, V196A, E238D, and P289L. CHC-B2:CHC-B1 ratios are pSE659 0.06:1; pSE657 0.06:1; pSE670 0.07:1, pSE622 0.07:1; pSE646 0.06:1; pSE639 0.07:1; pSE617 0.05:1; pSE682 0.05:1 and pSE621 0.06:1. *pSE617 was used for gene replacements

	1 PstI 100
pSE659	VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQGQGGSKDTGTTEVVYPVISVVCITTAAAWLFRRCLVERRLLFDALLFLGLLFTSWQSPLMNWFH
pSE657	VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQGQGGSKDTGTTEVVYPVISVVCITAAAAWLFRRCLVERRLLFDALLFLGLLFTSWQSPLMNWFH
pSE670	VMVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQGQGGSKDTGTTEVVYPVISVVCITAAAAWLFRRCLVERRLLFDALLFLGLLFTSWQSPLMNWFH
pSE622	VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQGQGGSKDTGTTEVVYPVISVVCITTAAAWLFRRCLVERRLLFDALLFLGLLFTSWQSPLMNWFH
pSE646	VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQGQGGSKDTGTTEVVYPVISVVCITAAAAWLFRRCLVERRLLFDALLFLGLLFTSWQSPLMNWFH
pSE639	VMVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQGQGGSKDTGTTEVVYPVISVVCITTAAAWLFRRCLVERRLLFDALLFLGLLFTSWQSPLMNWFH
pSE617*	VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQGQGGSKDTGTTEVVYPVISVVCITTAAAWLFRRCLVERRLLFDALLFLGLLFTSWQSPLMNWFH
pSE682	VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQGQGGSKDTGTTEVVYPVISVVCITTAAAWLFRRCLVERRLLFDALLFLGLLFTSWQSPLMNWFH
pSE621	VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGPGQGGSKDTGTTEVVYPVISVVCITTAAAWLFRRCLVERRLLFDALLFLGLLFTSWQSPLMNWFH
aveC WT	VVVWAGVGLLFLALQAYVFSRWAADGGYRLIETAGQGQGGSKDTGTTDVVYPVISVVCITAAAAWLFRRCRVERRLLFDALLFLGLLFASWQSPLMNWFH

	101 200
pSE659	SVLVSNASVWGAVGSWGPYVPGWQGAGPGAEAEMPPASASVCMSALIVSVLCSKALGWIKARRPAWRTWRLVLAVCFISIVLGLSEPLPSASGISAWARA
pSE657	SVLVSNASVWGAVGSWGPYVPGWQGAGPGAEAEMPPASASVCMSALIVSVLCSKALGWIKARRPAWRTWRLVLAVCFISIVLGLSEPLPSASGISAWARA
pSE670	SVLVSNASVWGAVGSWGPYVPGWQGAGPGAEAEMPPASASVCMSALIVSVLCSKALGWIKARRPAWRTWRLVLAVCFISTVLGLSEPLPSASGISAWARA
pSE622	SVLVSNASVWGAVGSWGPYVPGWQGAGPGAEAEMPLASTSVCMSALIVSVLCSKALGWIKARRPAWRTWRLVLAVCFISIVLGLSEPLPSASGISAWARA
pSE646	SVLVSNASVWGAVGSWGPYVPGWQGAGPGAEAEMPPASASVCMSALIVSVLCSKALGWIKARRPAWRTWRLVLAVCFISIVLGLSEPLPSASGISVWARA
pSE639	SVLVSNASVWGAVGSWGPYVPGWQGAGPGAEAEMPPASASVCMSALIVSVLCSKALGWIKARRPAWRTWRLVLAVCFISIVLGLSEPLPSASGISAWARA
pSE617*	SVLVSNASVWGAVGSWGPYVPGWQGAGPGAEAEMPPASASVCMSALIVSVLCSKALGWIKARRPAWRTWRLVLAVCFISIVLGLSEPLPSASGISAWARA
pSE682	SVLVSNASVWGAVGSWGPYVPGWQGAGPGAEAEMPPASASVCMSALIVSVLCSKALGWIKAHRPAWRTWRLVLAVCFISIVLGLSEPLPSASGISAWARA
pSE621	SVLVSNASVWGAVGSWGPYVPGWQGAGPGAEAEMPPASASVCMSALIVSVLCSKALGWIKARRPAWRTWRLVLAVCFISIVLGLSEPLPSASGISAWARA
aveC WT	SVLVSNASVWGAVGSWGPYVPGWQGAGPGAEAEMPLASASVCMSALIVTVLCSKALGWIKARRPAWRTWRLVLAVFFIGIVLGLSEPLPSASGISVWARA

	201 SphI 303
pSE659	LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDDRDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHILLSLVGGQPPDQLPDSFQAPAAY*
pSE657	LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDDRDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHVLLSLVGGQPPDQLPDSFQAPAAY*
pSE670	LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDDRDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHVLLSLVGGQPPDQLPDSFQAPAAY*
pSE622	LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDDRDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHVLLSLVGGQPPDQLPDSFQAPAAY*
pSE646	LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDDRDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHILLSLVGGQLPDQLPDSFQAPAAY*
pSE639	LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDDRDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHVLLSLVGGQPPDQLPDSFQAPAAY*
pSE617*	LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDDRDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHVLLSLVGGQPPDQLPDSFQAPAAY*
pSE682	LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDDRDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHVLLSLVGGQPPDQLPDSFQAPAAY*
pSE621	LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDDRDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHVLLSLVGGQPPDQLPDSFQAPAAY*
aveC WT	LPEVTLWSGEWYQFPVYQAVGSGLVCCMLGSLRFFRDERDESWVERGAWRLPQRAANWARFLAVVGGVNAVMFLYTCFHILLSLVGGQPPDQLPDSFQAPAAY*

transformed into protoplasts of *S. avermitilis* SE247-11 and analyzed by fermentation for avermectin production. Fermentation analysis showed that the pSE375 transformants were producing normal quantities of avermectins at a CHC-B2:CHC-B1 ratio of 0.07:1. To incorporate mutations conferring improved product ratio into the chromosome, single colony pSE375 transformants of *S. avermitilis* SE247-11 were screened for gene replacement by monitoring the loss of thiostrepton and erythromycin resistance (Stutzman-Engwall et al., 2003). Fermentation analysis showed that Erm^S $Thio^S$ transformants produced avermectins at a 0.07:1 CHC-B2:CHC-B1 ratio (Fig. 6). Gene replacement was verified by PCR amplification of the *aveC* gene and DNA sequence analysis. No ErmE, Thio or pWHM3 sequences were detected after PCR amplification using appropriate primers (data not shown), confirming the mutated *aveC* gene integrated by double cross-over.

4. Discussion

The ability to carry out molecular genetic manipulations on high titer production strains used in industrial fermentations is an extremely desirable but highly challenging objective. Here, we present an important milestone toward the goal of engineering the biosynthetic pathway of a natural product (avermectin) for pharmaceutical applications. By integrating DNA shuffling with HTP electrospray mass spectrometry, we developed a sensitive and accurate method to monitor polyketide production and to manipulate the polyketide synthetic pathway while maintaining overall avermectin expression in the production organism. This combination of technologies allows for the creation, modification and production of polyketides in ways that are inaccessible to traditional medicinal chemistry.

Directed evolution has been highly successful for proteins where simple and rapid HTP screens or selections are available. No such tool is generally available for polyketide screening. In addition, a major limitation for HTP screening of small molecule production from *Streptomyces* has been to achieve reproducible cell growth and production levels of independent variants of secondary metabolites. These challenges were overcome by improving the quality of library diversity, thereby reducing the number of variants that needed to be screened, and by establishing a robust 96 well solid phase growth and production screen. We describe a semi-synthetic shuffling format in which diversity is identified by sequencing and subsequently incorporated by oligonucleotides into the shuffling reaction. Semi-synthetic shuffling enables each combination of amino acid substitutions from a single gene shuffled library to occur at an equal probability. This strategy proved effective in generating high quality libraries, ultimately generating clones with a 23-fold improved activity over the wild-type gene, and containing an average of 10 amino acid substitutions. This was achieved after four rounds of directed evolution and screening of 28,000 independent clones in total.

While we succeeded in creating a new *S. avermitilis* strain having an improved CHC-B2:CHC-B1 ratio, the function of the AveC protein in the polyketide synthesis production pathway remains unclear. Although a functional or structural understanding of an underlying biological system can assist in its improvement, protein optimization through DNA shuffling does not depend on it. Instead, as in this example, we only required sequence-function relationship: the information that amino acid substitutions in *aveC* change the

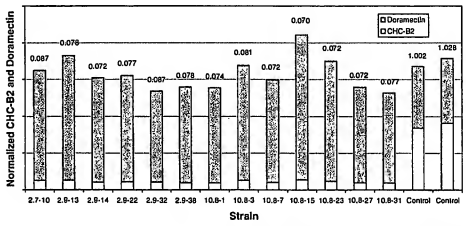


Fig. 6. Fermentation analysis of thirteen independent *S. avermitilis* transformants containing the mutated *aveC* allele (D48E, A61T, R71L, A89T, L136P, T149S, F176C, G179S, V196A, and E238D) introduced by gene replacement into the chromosome of the *S. avermitilis* production strain (normalized data derived from μg/ml production levels). The variant *aveC* gene directs the production of CHC-B2:CHC-B1 at a 0.07:1 ratio. The control strains produce CHC-B2:CHC-B1 at a 1:1 ratio.

CHC-B2:CHC-B1 ratio. By screening for an improved CHC-B2:CHC-B1 ratio and recombining the identified diversity, we have engineered a highly complex natural product pathway. The same approach can be used for other biological production pathways as long as a reproducible format for HTP growth and compound production is integrated with accurate analytical screens and DNA shuffling.

Acknowledgments

We thank Lu Liu, Jon Ness, Yong Hong Chen and Jeremy Minshull (Maxygen) Rick Davies and Jim Lengen (Pfizer) and Matt Tobin and Lori Giver (Codexis) for helpful discussions, expert technical assistance and critical reading of the manuscript.

References

Aita, T., Hamamatsu, N., Nomiya, Y., Uchiyama, H., Husimi, Y., 2002. Surveying a local fitness landscape of a protein with epistatic sites for the study of directed evolution. Biopolymers 64, 95–105.

Burg, R.W., Miller, B.M., Baker, E.E., Birnbaum, J., Currie, S.A., Hartman, R., Kong, Y.L., Monaghan, R.L., Olson, G., Putter, I., Tunac, J.B., Wallick, H., Stapley, E.O., Oiwa, R., Omura, S., 1979. Avermectins, new family of potent anthelmintic agents: producing organism and fermentation. Antimicrob. Agents Chemother. 15, 361–367.

Cane, D.E., Liang, T.C., Kaplan, L.K., Nullin, M.K., Schulman, M.D., Hensens, O.D., Douglas, A.W., Albers-Schonberg, G., 1983. Biosynthetic origin of the carbon skeleton and oxygen atoms of the avermectins. J. Am. Chem. Soc. 105, 4110–4112.

Chartrain, M., Salmon, P.M., Robinson, D.K., Buckland, B.C., 2000. Metabolic engineering and directed evolution for the production of pharmaceuticals. Curr. Opin. Biotechnol. 11, 209–214.

Chen, T.S., Arison, B.H., Gullo, V.P., Inamine, E.S., 1989. Further studies on the biosynthesis of the avermectins. J. Ind. Microbiol. 4, 231–238.

Chen, T.S., Inamine, E.S., 1989. Studies on the biosynthesis of avermectins. Arch. Biochem. Biophys. 270, 521–525.

Crameri, A., Stemmer, W.P.C., Minshull, J., Bass, S.H., Welch, M., Ness, J.E., Gustafsson, C., Patten, P.A., 2000. Oligonucleotide mediated nucleic acid recombination. PCT International Patent Application WO 00/42561.

Devereux, J., Haeberli, P., Smithies, O., 1984. A comprehensive set of sequence analysis programs for the VAX. Nucleic Acids Res. 12, 387–395.

Dutton, C.J., Gibson, S.P., Goudie, A.C., Holdom, K.S., Pacey, M.C., Ruddock, J.C., Bu'Lock, J.D., Richards, M.S., 1991. Novel avermectins produced by mutntional biosynthesis. J. Antibiot. 44, 357–365.

Goudie, A.C., Evans, N.A., Gration, K.A.F., Bishop, B.F., Gibson, S.P., Holdom, K.S., Kaye, B., Wicks, S.R., Lewis, D., Weatherley, A.J., Bruce, C.I., Herbert, A., Seymour, D.J., 1993. Doramectin—a potent novel endectocide. Vet. Parasitol. 49, 5–15.

Hafner, E.W., Holley, B.W., Holdom, K.S., Lee, S.E., Wax, R.G., Beck, D., McArthur, H.A.I., Wernau, W.C., 1991. Branched-chain fatty acid requirement for avermectin production by a mutant of *Streptomyces avermitilis* lacking branched-chain 2-oxo acid dehydrogenase activity. J. Antibiot. 44, 349–356.

Ikeda, H., Nonomiya, T., Usami, M., Ohta, T., Omura, S., 1999. Organization of the biosynthetic gene cluster for the polyketide anthelmintic macrolide avermectin in *Streptomyces avermitilis*. Proc. Natl. Acad. Sci. USA 96, 9509–9514.

Ikeda, H., Omura, S., 1995. Control of avermectin biosynthesis in *Streptomyces avermitilis* for the selective production of a useful component. J. Antibiot. 48, 549–562.

Ikeda, H., Pang, C.-H., Endo, H., Ohta, T., Tanaka, H., Omura, S., 1995. Construction of a single component producer from the wild type avermectin producer *Streptomyces avermitilis*. J. Antibiot. 48, 532–534.

Katz, L., Donadio, S., 1993. Polyketide synthesis: prospects for hybrid antibiotics. Ann. Rev. Microbiol. 47, 875–912.

Kieser, T., Bibb, M.J., Buttner, M.J., Chater, K.F., Hopwood, D.A., 2000. Practical Streptomyces Genetics. John Innes Foundation, Norwich.

MacNeil, D.J., Occi, J.L., Gewain, K.M., MacNeil, T., Gibbons, P.H., Ruby, C.L., Danis, S.J., 1992. Complex organization of the *Streptomyces avermitilis* genes encoding the avermectin polyketide synthase. Gene 115, 119–125.

Moore, C.J., Jin, H-M., Kuchner, O., Arnold, F.A., 1997. Strategies for the in vitro evolution of protein function: enzyme evolution by random recombination of improved sequences. J. Mol. Biol. 272, 336–347.

Patrick, W.M., Firth, A.E., Blackburn, J.M., 2003. User-friendly algorithms for estimating the completeness and diversity in randomized protein-encoding libraries. Protein Eng. 16, 451–457.

Pacey, M.S., Dutton, C.J., Monday, R.A., Ruddock, J.C., Smith, G.C., 2000. Preparation of 13-*epi*-Selamectin by biotransformation using a blocked mutant of *Streptomyces avermitilis*. J. Antibiot. 53, 301–305.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Stemmer, W.P.C., 1994a. Rapid evolution of a protein in vitro by DNA shuffling. Nature 370, 389–391.

Stemmer, W.P.S., 1994b. DNA in vitro shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. Proc. Natl. Acad. Sci. USA 91, 10747–10751.

Stutzman-Engwall, K.J., Conlon, S.W., Fedechko, R.W., Kaczmarek, F.S., 1997. Engineering *Streptomyces avermitilis* biosynthetic genes. In: Hutchinson, C.R., McAlpine, J. (Eds.), Developments in Industrial Microbiology Series, vol. 35. Society for Industrial Microbiology, Fairfax, VA, pp. 7–13.

Stutzman-Engwall, K.J., Conlon, S., Fedechko, R., Kaczmarek, F., McArthur, H., Krebber, A., Chen, Y., Minshull, J., Raillard, S.A., Gustafsson, C., 2003. Engineering the *aveC* gene to enhance the ratio of doramectin to its CHC-B2 analogue produced in *Streptomyces avermitilis*. Biotech. Bioeng. 82, 359–369.

Stutzman-Engwall, K.J., McArthur, H., Katoh, Y., 2001. *Streptomyces avermitilis* gene directing the ratio of B2:B1 avermectins. US Patent 6,248,579.

Vajdos, F.F., Adams, C.W., Breece, T.N., Presta, L.G., de Vos, A.M., Sidhu, S.S., 2002. Comprehensive functional maps of the antigen-binding site of an anti-ErbB2 antibody obtained with shotgun scanning mutagenesis. J. Mol. Biol. 320 (2), 415–428.

Varu, J., Lewandowska-Skarbek, M., Wang, Y.-G., Donadio, S., Hutchinson, C.R., 1989. Cloning of genes governing the deoxysugar portion of the erythromycin biosynthetic pathway in *Saccharopolyspora erythraea* (*Streptomyces erythreus*). J. Bacteriol. 171, 5872–5881.

Appl Microbiol Biotechnol (2005) 68: 228–236
DOI 10.1007/s00253-005-1893-6

Jörg Mampel · Hartwig Schröder · Stefan Haefner ·
Uwe Sauer

Single-gene knockout of a novel regulatory element confers ethionine resistance and elevates methionine production in *Corynebacterium glutamicum*

Received: 29 October 2004 / Revised: 17 December 2004 / Accepted: 19 December 2004 / Published online: 25 January 2005

Abstract Despite the availability of genome data and recent advances in methionine regulation in *Corynebacterium glutamicum*, sulfur metabolism and its underlying molecular mechanisms are still poorly characterized in this organism. Here, we describe the identification of an ORF coding for a putative regulatory protein that controls the expression of genes involved in sulfur reduction dependent on extracellular methionine levels. *C. glutamicum* was randomly mutagenized by transposon mutagenesis and 7,000 mutants were screened for rapid growth on agar plates containing the methionine antimetabolite D,L-ethionine. In all obtained mutants, the site of insertion was located in the ORF NCgl2640 of unknown function that has several homologues in other bacteria. All mutants exhibited similar ethionine resistance and this phenotype could be transferred to another strain by the defined deletion of the NCgl2640 gene. Moreover, inactivation of NCgl2640 resulted in significantly increased methionine production. Using promoter *lacZ*-fusions of genes involved in sulfur metabolism, we demonstrated the relief of L-methionine repression in the NCgl2640 mutant for cysteine synthase, *o*-acetylhomoserine sulfhydrolase (*metY*) and sulfite reductase. Complementation of the mutant strain with plasmid-borne NCgl2640 restored the wild-type phenotype for *metY* and sulfite reductase.

J. Mampel (✉) · U. Sauer
Institute of Biotechnology,
Swiss Federal Institute of Technology (ETH),
Zurich, 8093, Switzerland
e-mail: mampel@micro.biol.ethz.ch
Tel.: +41-1-6323849
Fax: +41-1-6321378

J. Mampel
Institute of Microbiology,
Swiss Federal Institute of Technology (ETH),
Zurich, 8093, Switzerland

H. Schröder · S. Haefner
BASF Aktiengesellschaft,
Ludwigshafen, 67056, Germany

Introduction

The two sulfur-containing amino acids methionine and cysteine are key elements in animal feed. While methionine is an essential amino acid, animals are able to convert supplemented methionine into cysteine. Both amino acids are not only essential for protein biosynthesis, but are also precursors of various metabolites such as glutathione, *S*-adenosylmethionine, polyamines and biotin, and are involved as the methyl group donor in numerous cellular processes. Since concentrations of methionine and cysteine are often low in edible plant sources (Nikiforova et al. 2002), methionine is an important feed additive, with an annual production in the range of 300,000 t, third to glutamate and lysine (Leuchtenberger 1996). Unlike other bulk amino acids, methionine is produced by chemical synthesis and thus is the last of the major commercial amino acids that is not produced by fermentation.

Corynebacterium glutamicum strains have proven to be effective producers of L-lysine and L-threonine (Sahm et al. 1996; Hermann 2003) which, together with L-methionine, belong to the aspartate family of amino acids. Although lysine titers of 120 g/l or more are routinely obtained in industrial fermentations with *C. glutamicum*, no such process exists for methionine. The key step that distinguishes the biosynthesis of lysine and methionine is the incorporation of sulfur into the carbon skeleton. The common source of sulfur is sulfate, that has to be taken up, activated and reduced by the consumption of 7 mol ATP and 8 mol NADPH per mole of methionine (Neidhardt et al. 1990). In terms of cellular energy demands, this makes methionine the most expensive amino acid.

Genes involved in methionine biosynthesis are scattered on the genome of *C. glutamicum* (Ruckert et al. 2003), while genes involved in assimilatory sulfate reduction are clustered and at least partially organized in the *cys*-operon (Fig. 1). Sulfite reductase (NCgl2718) generates sulfide and is one of the genes under the control of the putative *cys*-operon promoter. Sulfide is then incorporated into *o*-acetylhomoserine via direct sulfhydrylation (MetY), or into *o*-acetylserine to give cysteine via the transsulfhydrylation

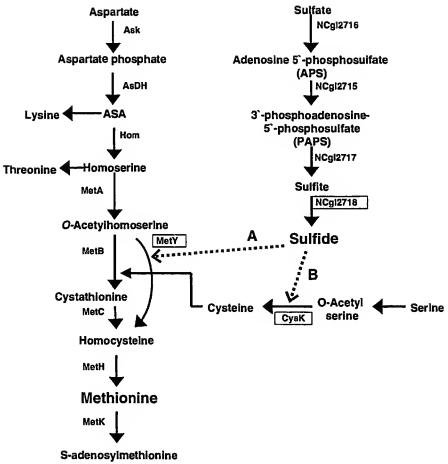


Fig. 1 The split pathway of sulfur incorporation in methionine biosynthesis. Sulfide is incorporated either via direct sulfhydrolation that is catalyzed by MetY (a), or by the trans-sulfhydrolation pathway (b). *lacZ* fusions were generated for promoters of the boxed gene products. The promoter of NCgl2718 governs the expression of genes involved in assimilatory sulfate reduction organized in a gene cluster (*cys*-operon, NCgl2715–NCgl2720). Cysteine synthase (NCgl2473) is not part of this cluster. *Ask* Aspartate kinase, *AsDH* aspartate semialdehyde dehydrogenase, *CysK* cysteine synthase, *Hom* homoserine dehydrogenase, *MetA* homoserine acetyltransferase, *MetB* cystathionine γ-synthase, *MetC* cystathionine β-lyase, *MetH* methionine synthase, *MetY O*-acetylhomoserine sulfhydrolase, *MetK S*-adenosylmethionine synthase, *NCgl2715* sulfate-adenosyltransferase subunit 1, *NCgl2716* sulfate-adenosyltransferase subunit 2, *NCgl2717* PAPS-reductase, *NCgl2718* sulfite reductase annotated as putative nitrite reductase

pathway (cysteine synthase CysK, NCgl2473; Fig. 1). *C. glutamicum* and related organisms use both pathways for methionine biosynthesis (Hwang et al. 2002; Lee and Hwang 2003). A diversity of regulatory mechanisms controlling the metabolic flux through these pathways has been described (Lee and Hwang 2003). Tight regulation avoids uneconomic depletion of cellular energy, but deregulation is indispensable for the overproduction of methionine. A detailed understanding of the regulatory mechanisms involved in sulfhydrylation is thus essential for the future rational design of methionine-producing strains.

There are only scattered reports on methionine-producing variants of *C. glutamicum*. Since 1975, when Kase and Nakayama (1975) reported on a multistep procedure using random mutagenesis and antimetabolite selection, no further improvement has been reported. Recently, a putative transcriptional repressor (McbR) involved in the regulation of the metabolic network directing the synthesis of sulfur containing amino acids was identified and a *mcbR* knockout strain was constructed (Rey et al. 2003). Still, methionine levels obtained in both examples fall short of commercial relevance.

Here, we report the identification of a novel regulator of methionine biosynthesis by screening a transposon library of *C. glutamicum* for ethionine-resistant strains. As a structural analogue of methionine, D,L-ethionine cannot be metabolized and thus mimics high concentrations of methionine. Consequently, methionine biosynthesis is down-regulated and the organism finally starves from methionine depletion. One of several mechanisms to circumvent antimetabolite toxicity is to outcompete the toxic agent by overproduction of the natural metabolite (Mondal and Chatterjee 1994). Overproducing strains are typically based on mutations that relieve the feed-back inhibition of biosynthetic proteins (Kase and Nakayama 1975). Here, we screened for overproduction upon Tn*5531*-inactivation of genes involved in regulatory cascades.

Materials and methods

Bacterial strains, media and plasmids

C. glutamicum ATCC14752 or ATCC13032 were routinely cultivated in CGXII minimal medium (Keilhauer et al. 1993). *Escherichia coli* DH5α was used for standard cloning and *E. coli* ET12567 for plasmid amplification when plasmids were destined to be transformed in *C. glutamicum*. Strains and plasmids used in this study are listed in Table 1. Luria broth (LB) supplemented with appropriate antibiotics (50 μg/ml kanamycin, 50 μg/ml

Table 1 Strains and plasmids used in this study

Strain or plasmid	Relevant phenotype	Source or reference
C. glutamicum strains		
ATCC13032	Wild type	ATCC
ATCC14752	Wild type	ATCC
13032::2640	Wild type with *NCgl2640*, Km^r	This work
14752::2640	Wild type with *NCgl2640*, Km^r	This work
14752-PcysK	ATCC14752 with reporter plasmid pClik-P*cysK*, Cm^r	This work
CG::2640-PcysK	14752::2640 with reporter plasmid pClik-P*cysK*, Km^r Cm^r	This work
CG::2640-PcysK-cpl	14752::2640 with reporter plasmid pClik-P*cysK*-cpl, Km^r Cm^r	This work
14752-PmetY	ATCC14752 with reporter plasmid pClik-P*metY*, Cm^r	This work
CG::2640-PmetY	14752::2640 with reporter plasmid pClik-P*metY*, Km^r Cm^r	This work
CG::2640-PmetY-cpl	14752::2640 with reporter plasmid pClik-P*metY*-cpl, Km^r Cm^r	This work
14752-Pcys	ATCC14752 with reporter plasmid pClik-Pcys, Cm^r	This work
CG::2640-Pcys	14752::2640 with reporter plasmid pClik-Pcys, Km^r Cm^r	This work
CG::2640-Pcys-cpl	14752::2640 with reporter plasmid pClik-Pcys-cpl, Km^r Cm^r	This work
E. coli strains		
DH5α	F^-*endA1*, *hsdR17*(rk^-mk^+) *supE44*, *thi-1* λ^-*recA1 gyrA96 relA1* φ*80*Δ*lacΔm15*	Hanahan (1983)
ET12567	*dam dcm hsd*, restriction deficient	MacNeil et al. (1992)
Plasmids		
pClik	*E. coli–C. glutamicum* shuttle vector; replicative in *C. glutamicum*, medium copy (10–20) number; Cm^r	EP 03/05423
pUC18	Ap^r, *lacZ*	Stratagene
pCGL0040	Donor of Tn*5531* (IS*1207* Km^r); Ap^r, $ori V_{E.c.}$	Ankri et al. (1996b)
pClik-SacB	Vector for allelic exchange by homologous recombination; nonreplicative in *C. glutamicum* Km^r, SacB	EP 03/05423
pSdel-NCgl2640	pClikSacB-based allelic exchange vector for chromosomal deletion of NCgl2640	This work
pClik-PcysK	*cysK*-promoter-probe vector *lacZ*-fusion, Cm^r	This work
pClik-PcysK-cpl	pClik-P*cysK* with NCgl2640, Cm^r	This work
pClik-PmetY	*metY*-promoter-probe vector *lacZ*-fusion, Cm^r	This work
pClik-PmetY-cpl	pClik-P*metY* with NCgl2640, Cm^r	This work
pClik-Pcys	*cys*-operon-promoter-probe vector *lacZ*-fusion, Cm^r	This work
pClik-Pcys-cpl	pClik-Pcys with NCgl2640, Cm^r	This work

chloramphenicol, 100 μg/ml ampicillin) was the standard medium for *E. coli* strains. LB medium supplemented with 4 mM $MgSO_4$ and 10 mM KCl (Psi-broth) was the recovery medium for chemically transformed *E. coli*. The recovery medium for electroporated *C. glutamicum* strains was LB with brain heart infusion and sorbitol (LBHIS; Liebl et al. 1989). Plasmid pCGL0040 (GenBank accession no. U53587) was used as the donor for Tn*5531* (IS*1207* Km^r) and was amplified in *E. coli* ET12567.

Recombinant DNA technologies

The transformation of *E. coli* cells with plasmid DNA was performed with chemically competent *E. coli* DH5α or ET12567. Cells were prepared according to the rubidium chloride method (http://micro.nwfsc.noaa.gov/protocols/) and transformed as described by Sambrook et al. (1989). Preparation of competent *C. glutamicum* cells and electrotransformation was done as described elsewhere (Liebl et al. 1989; Ankri et al. 1996a).

Genomic DNA was isolated using the Wizard genomic DNA purification kit (Promega). Plasmid preparation from *E. coli* cells was routinely done with the QIAprep miniprep kit (Qiagen). Restriction endonucleases were from Roche Diagnostics. Digested DNA fragments were recovered from agarose gels with the QIAEX II gel extraction kit (Qiagen). Standard DNA techniques were performed as described by Sambrook et al. (1989). DNA sequencing was performed using the Global edition IR2 system (LI-COR, Lincoln, Neb.).

The genome database used to identify ORFs and promoters was ERGO (Integrated Genomics, Chicago, Ill.). The NCgl numbers refer to the *C. glutamicum* ATCC13032 genome sequence in the GenBank database at the National Center for Biotechnology Information (NCBI; GenBank accession no. NC_003450). The promoters deduced from the ERGO database for cysteine

synthase *cysK* (NCgl2473; promoter position 2721625–2721822), sulfite reductase (NCgl2718; promoter position 3005188–3005389) and *o*-acetylhomoserine-sulfhydrolase (*metY*, NCgl0625; promoter position 667771–668107) were PCR-amplified and fused by *Xho*I and *Bam*HI linkers to promoterless *lacZ* on plasmid pClik (Cm^r), opened by the same restriction enzymes. The resulting plasmids were termed pClik-PcysK, pClik-Pcys and pClik-PmetY, respectively (Table 1). A unique *BglII* site on these plasmids was used to introduce PCR-amplified NCgl2640, using primer pair 2640-fwd-BglII (5′-CCGCTGCTGCT GGTGGCGCTAGATCTGCTAACGGC-3′) and 2640-rev-BglII (5′-ATGTGTTGGGAGATCTCTTAAGTTATTTAG TCCAG-3′). The amplified DNA fragment comprised putative regulatory elements located up to 370 bp upstream of the NCgl2640 gene.

Transposon mutagenesis, screening and localization of transposon insertion sites

Plasmid pCGL0040 was isolated from *E. coli* ET12567 and *C. glutamicum* ATCC14752 was transformed with the plasmid by electroporation. Transposon insertion mutants were selected by plating on LBHIS containing 20 μg/ml kanamycin. All mutants were pooled, washed twice with sterile 0.9% NaCl and plated on CGXII containing kanamycin (20 μg/ml) and ethionine (7.5 g/l) in 100-μl aliquots of a 10^6 dilution of the pool. The most rapidly growing colonies were selected for further analysis.

For the localization of the transposon insertion sites, genomic DNA of the mutants was isolated and digested with *Eco*RI, as described by Simic et al. (2001). Insertion sites were determined by cloning transposon–chromosome junction sites into *Eco*RI-digested pUC18. Plasmids from kanamycin-resistant clones were isolated and subsequently sequenced with oligonucleotide Tn5531-Eco (5′-CGGG TCTACACCGCTAGCCCAGG-3′; Simic et al. 2001). The sequences thus obtained were analyzed using the BLASTn program applied to the NCBI GenBank sequence and the ERGO database. Different sequence analysis tools (http://www.expasy.org/, http://npsa-pbil.ibcp.fr or the protein family database, PFAM at http://pfam.wustl.edu/) were used for pattern and profile searches with the NCgl2640 sequence.

Chromosomal deletion of NCgl2640 in *C. glutamicum* ATCC13032

Using the two primer pairs 2640-SacB1 (5′-GAGAGGGC CCATCAGCAGAACCTGGAACC-3′) with 2640-SacB2 (5′-GATCCAGAGGTCCACAACC-3′) and 2640-SacB3 (5′-GATGGTTCAAGACGAACTCC-3′) with 2640-SacB4 (5′-GAGAGTCGACCAGAATCAATTCCAGCCTTC-3′), the upstream and downstream region of NCgl2640 was PCR-amplified from chromosomal DNA of *C. glutamicum* ATCC13032. The resulting fragments were digested with *ApaI*/*Xba*I and *Spe*I/*Sal*I, respectively, and cloned together into the nonreplicative vector pClik-SacB, that was digested with *ApaI*/*SalI*, yielding plasmid pSdel-NCgl2640. *C. glutamicum* ATCC13032 was transformed by electroporation with the nonreplicative plasmid pSdel-NCgl2640. Kanamycin-resistant clones contained chromosomally integrated plasmids. Subsequently, we selected for loss of the plasmid by screening for sucrose-resistant mutants according to Schaefer et al. (1994). The deletion was verified by PCR analysis and Southern blotting.

LacZ activity measurements

Selected mutant strains of *C. glutamicum* were transformed with the plasmids pClik-PcysK, pClik-PmetY and pClik-Pcys or the NCgl2640-complemented derivatives thereof (Table 1). Transformants were grown in CGXII minimal medium containing chloramphenicol (15 μg/ml) in the presence or absence of 10 mM L-methionine. Cells were grown to an optical density at 600 nm (OD_{600}) of 1–3 and analyzed for β-Gal activity, as described by Sambrook et al. (1989). Assays were done in triplicates in four independent test series.

Isolation of DNA-binding proteins

The principle of isolating DNA-binding proteins by DNA-affinity chromatography using magnetic beads is essentially described by Gabrielsen and Huet (1993) and a detailed protocol for *C. glutamicum* is available (Rey et al. 2003). We basically followed the latter protocol with a few exceptions. All buffers except the elution buffers were supplemented with 2.5 mM L-methionine. Crude extracts of cells grown in the presence or absence of 10 mM L-methionine were prepared separately and combined thereafter. Immediately after cell disruption, crude extract was protected against proteolysis by a protease inhibitor cocktail (phenylmethylsulfonyl fluoride, aprotinin, leupeptin; Rosenberg 1996). After ultracentrifugation of the crude extract (200,000 *g*, 40 min, 4°C), the protein solution was desalted by gel filtration (Sephadex G25). Biotinylated PCR-amplified promoter-DNA was immobilized on streptavidin-coated Dynabeads (M270; Dynal Biotech). As a negative control fragment, we amplified a 460-bp fragment from the upstream region of the *groES* gene of *C. glutamicum*. The washing buffer contained high amounts of unspecific competitor DNA (0.4 mg/ml salmon testes DNA; Sigma). 1D-SDS-PAGE was done with a 4% stacking gel and a 12% separative gel (Schägger and von Jagow 1987) and proteins were stained with colloidal Coomassie brilliant blue G-250 (Neuhoff et al. 1988). The protocol for tryptic digestion and matrix assisted laser desorption/ionization–time of flight (MALDI-TOF) analysis was essentially that described by Hermann et al. (2001).

Determination of extra- and intracellular methionine concentrations

Methionine was quantified as its *o*-phthaldialdehyde derivative by high-pressure liquid chromatography (HPLC; Molnar-Perl 2001). *C. glutamicum* was grown to the stationary phase in 500-ml shake-flasks with culture volumes of 50–100 ml at 30°C and 225 rpm. Cells were removed by centrifugation (10,000 g, 10 min, 4°C) and methionine concentrations were determined by HPLC analysis.

For the determination of intracellular methionine levels, cells were separated from the bulk liquid and inactivated by silica oil centrifugation, as described by Ebbighausen et al. (1989), and subsequently disrupted by sonification or by a blue-capped Ribolyser (FastPrep; Q-Biogene). Soluble protein in the supernatant was measured by a Bradford-type assay (Bradford 1976). The intracellular content of soluble protein in *C. glutamicum* was empirically determined as 250 mg/ml. Based on this value, we calculated the total internal cell volume of a sample in order to determine methionine concentrations upon HPLC analysis.

Results

Selection and identification of ethionine-tolerant transposon mutants

We hypothesized that ethionine tolerance can be acquired by overproduction of methionine, which in turn can be achieved by inactivation of a repressor involved in the regulation of methionine biosynthesis. Therefore, we transformed *C. glutamicum* ATCC14752 with pCGL0040 as the donor of Tn*5531* by electroporation. About 7,000 mutants were obtained on LBHIS plates containing kanamycin. All mutants were scraped off the plates, pooled, washed and plated onto CGXII-agar plates containing 7.5 g/l D,L-ethionine plus kanamycin. We plated about 100,000 colony-forming units (CFU) to ensure multiple recoveries of putatively ethionine-resistant mutants. Growth of the wild type was inhibited by 6 g/l ethionine for at least 4 days. After 2 days, 11 kanamycin- and ethionine-resistant mutants were isolated. All mutants contained the identical Tn*5531* insertion in ORF NCgl2640. This mutation was termed 14752::2640; and one clone was selected for further experimentation (Table 1).

Sequence analysis of transposon insertion loci

The site of transposon insertion in strain 14752::2640 was located in the C-terminal half of the putative protein at position 2,918,026/2,918,027 (GenBank accession no. NC_003450). NCgl2640 was separated from NCgl2639 by 7 bp, so both genes are presumably organized as an operon (Fig. 2). NCgl2639 is annotated as a putative hydrolase or acetyltransferase (GenBank). NCgl2640 encodes a protein of 42 kDa and homology searches identified more than 25 putative bacterial proteins of significant homology (*e*-value

Fig. 2 Genomic context of NCgl2640. GenBank annotations available for the open reading frames: *NCgl2638* similarities to multisubunit Na^+/H^+ antiporter, *NCgl2639* similar to predicted hydrolases or acyltransferases, alpha/beta hydrolase superfamily, *NCgl2640* hypothetical protein/uncharacterized BCR, *NCgl2641* hypothetical protein, no annotation. *Tn5531* Transposon 5531 (*IS1207*). *Filled triangle* Site of insertion of Tn*5531* in NCgl2640

<2e–20), none of which were functionally assigned. A conserved domain search identified a consensus pattern for proteins of unknown function (COG2170) with high significance and for motif 04107 of the PFAM database that is characteristic for the glutamate-cysteine ligase family. No further consensus patterns were detected, in particular none for DNA-binding proteins.

Verification of the ethionine-resistance phenotype

C. glutamicum ATCC14752 and the mutant strain were cultivated in shake-flasks in CGXII medium containing 3 g/l glucose in the presence or absence of 7.5 g/l D,L-ethionine. Growth of the mutant in the presence of ethionine was indistinguishable from growth without ethionine or growth of the wild type without ethionine (Fig. 3).

To exclude second-site mutations that may have caused the ethionine-resistance phenotype, we tested the effect of NCgl2640 inactivation in *C. glutamicum* ATCC13032. NCgl2640 was excised by homologous recombination and selection of sucrose-tolerant mutant strains. This mu-

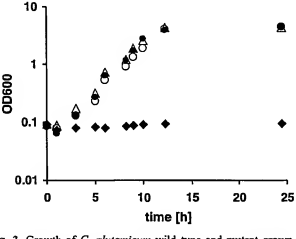


Fig. 3 Growth of *C. glutamicum* wild type and mutant grown in CGXII medium in the presence or absence of ethionine. The glucose concentration was 3 g/l, the D,L-ethionine concentrations was 7.5 g/l. *Filled symbols* Ethionine present during cultivation, *open symbols* cultivation without ethionine. *Filled triangles* Wild type, *filled circles* mutant strain

tant, designated 13032::2640, was readily resistant to D,L-ethionine at 7.5 g/l, while no growth of wild-type ATCC13032 was detected on these plates. Thus, we demonstrated that the ethionine-resistance phenotype depended on inactivation of NCgl2640.

Increased L-methionine levels in the NCgl2640 mutant

The resistance of *C. glutamicum* 14752::2640 to high ethionine concentrations may be due to upregulated biosynthesis of L-methionine. To test this hypothesis, we analyzed the production of L-methionine in the wild type and mutant in CGXII minimal medium batch cultures. Generally, the mutant accumulated about twice the amount of methionine compared with the wild type, both intra- and extracellularly (Fig. 4).

Altered expression levels of methionine biosynthesis genes in the NCgl2640 mutant

We assumed that tight regulation of sulfur incorporation would be a key element in methionine biosynthesis. Hence, inactivation of a repressor of sulfur gene expression would cause elevated methionine levels that confer ethionine resistance. Therefore, we decided to elucidate the impact of the NCgl2640 knockout on the expression levels of *metY*, *cysK* and *cys*-operon genes. For this purpose, strain 14752::2640 and the wild type were transformed with the *lacZ*-reporter plasmids pClik-PcysK, pClik-PmetY and pClik-Pcys (Table 1) and grown to an OD_{600} of 3.0 in CGXII medium with or without 10 mM L-methionine. Gene expression was then detected by monitoring LacZ activity.

In the wild type, the presence of methionine reduced the expression levels of all examined genes, but expression of the *cys*-operon was completely abolished. In the mutant strain, significant derepression was observed for the *cys*-operon (Fig. 5). A complementation assay was used to exclude the polar effects of the inactivation of NCgl2640 on the adjacent NCgl2639 (hydrolase or acetyltransferase; Fig. 2). Expression of NCgl2640 from medium-copy-number plasmids under the control of its endogenous promoter (370 bp of the upstream sequence were included) completely restored the wild-type phenotype in the mutant, i.e. methionine-induced repression of the *cys*-operon (Fig. 5).

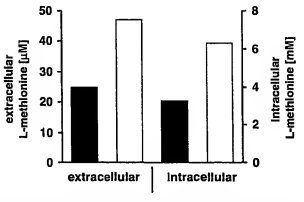


Fig. 4 Extra- and intracellular L-methionine levels in *C. glutamicum* wild type and the NCgl2640 knockout mutant in CGXII medium. *Dark bars C. glutamicum* wild type, *white bars* NCgl2640 knockout mutant

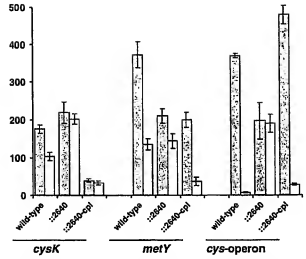


Fig. 5 Influence of the NCgl2640 knockout on methionine-dependent expression of cysteine synthase (*cysK*), *o*-acetylhomoserine sulfhydrolase (*metY*) and the *cys*-operon. *C. glutamicum* ATCC14752 (wild type), mutant strain 14752::2640 and mutant strains complemented with plasmid-borne NCgl2640 (::2640-cpl), harboring reporter plasmids pClik-P*cysK*, pClik-P*metY* or pClik-P*cys* were grown in CGXII medium (3 g/l glucose) in the presence (*white bars*) or absence (*dark bars*) of L-methionine (10 mM). Promoter activity was determined by LacZ activity reporter assays and quantified in Miller units

NCgl2640 does not bind to the *cysK*, *metY* and *cys*-operon promoters

We clearly demonstrated that expression of the *cys*-operon was modulated by NCgl2640. Since classic DNA-binding motifs could not be identified, we employed DNA-affinity purification in a pull-down assay to investigate whether NCgl2640 could bind to the respective promoter regions. PCR-amplified and bead-immobilized promoters were incubated in the presence of 2.5 mM L-methionine with combined crude extracts of *C. glutamicum* cells cultivated in the presence or absence of 10 mM L-methionine. Proteins that were eluted from the promoters at high salt concentrations (>200 mM) were separated by 1D-SDS-PAGE and analyzed by MALDI-TOF. Employing a similar approach, Rey et al. (2003) isolated the McbR-repressor along with four additional proteins that appeared to bind specifically to the *metY*-promoter. We largely confirmed these results in our study. In addition, we found that one of the proteins, exopolyphosphatase, reported by Rey et al. (2003) to bind specifically to the *metY*-promoter, also binds

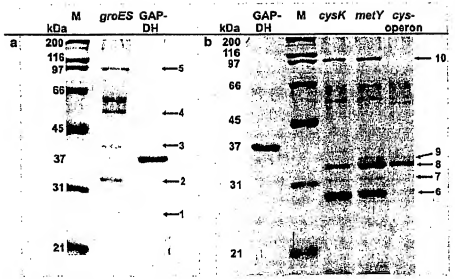


Fig. 6 SDS-PAGE of proteins binding to the putative control-promoter regions of *groES* (**a**) and to putative promoter regions of the *C. glutamicum* genes for cysteine synthase (*cysK*), *o*-acetylhomoserine sulfhydrolase (*metY*) and genes of the putative *cys*-operon (**b**). *1* Single-strand binding protein, *2* exopolyphosphatase, *3* diadenosine-tetraphosphate hydrolase, *4* DNA-polymerase III α-subunit, *5* DNA-polymerase I, *6* TetR-like regulator McbR that was shown to be involved in the regulation of methionine biosynthesis (Rey et al. 2003), *7* exopolyphosphatase (degradation product), *8* exopolyphosphatase, *9* ATP-phosphoribosyltransferase, *10* DNA-polymerase I. Except for exopolyphosphatase and DNA-polymerase I, none of the identified proteins bound to the *groES* control-promoter fragment (Fig. 6a). *GAP-DH* Glyceraldehyde-3-phosphate dehydrogenase used as additional marker-protein (37 kDa), *M* protein marker

to the control promoter of *groES*, indicating unspecific binding. Moreover, we could show that McbR also binds to promoters of *cysK* and the *cys*-operon; but binding to the *cys*-operon promoter was low compared to the *metY*-promoter (Fig. 6). Consistent in both studies, however, NCgl2640 was not detected, indicating that a direct DNA–protein interaction of NCgl2640 is not involved in the observed NCgl2640-mediated regulation.

Discussion

Using Tn*5531* mutagenesis and antimetabolite selection, we demonstrated that a single gene inactivation event renders *C. glutamicum* tolerant to high concentrations of ethionine and increases methionine biosynthesis. Classic strain development encompasses several rounds of random mutagenesis and antimetabolite selection (Sauer 2001); and the structural analogue of methionine, ethionine, is the major antimetabolite used for the selection of methionine overproducers (Lawrence et al. 1968; Kase and Nakayama 1975; Tani et al. 1988). Ethionine inhibition of methionine biosynthesis can be overcome by mutations that relieve feedback inhibition, e.g. by altering allosteric binding sites, but the resulting resistance is typically based on combinations of unidentified mutations. Here, we demonstrate that inactivation of NCgl2640 is sufficient to confer high-level tolerance to ethionine. To the best of our knowledge, this is the first report on a single-gene knockout that confers ethionine resistance.

Doubling of the methionine pool apparently suffices to confer resistance to a roughly 10-fold excess of the antimetabolite. The observed increase in extracellular methionine levels equals that reported recently for the *mcbR* mutant of *C. glutamicum* (Rey et al. 2002). Thus, inactivation of the putative master regulator of methionine biosynthesis has a similar effect on methionine production as the inactivation of NCgl2640, which underlines the importance of additional regulatory mechanisms in methionine and sulfur metabolism.

Based on *lacZ*-promoter assays, NCgl2640 was shown to exhibit a regulatory function in methionine biosynthesis that is most pronounced for genes involved in sulfur metabolism (*cys*-operon). In contrast to the putative transcriptional repressor McbR, whose activity does not appear to be directly induced by methionine, NCgl2640-based regulation is clearly dependent on extracellular methionine. Two arguments render a direct transcriptional regulation by NCgl2640 unlikely. First, no DNA-binding motif was detected, and second, we isolated McbR but not NCgl2640 with the DNA-affinity purification assay. Thus, NCgl2640 is probably an indirectly acting regulatory element involved in the repression of sulfur metabolism genes.

Two hypotheses on the actual regulation mechanism of NCgl2640 can be formulated from its annotation in the genome sequence of *C. glutamicum*. In the first, NCgl2640 might modulate gene expression indirectly via the putative master regulator McbR that was shown to bind directly to the *metY*-promoter (Rey et al. 2003). Based on one annotation of NCgl2640 as a putative glutamyl-cysteine-ligase, it might be speculated that interference with glutathione biosynthesis might affect McbR directly or indirectly, resulting in altered expression levels of McbR-

controlled genes. We observed a less pronounced effect of the inactivation of NCgl2640 on *cysK* and *metY* expression levels compared with *cys*-operon. By a proteomic analysis, Rey et al. (2003) showed that the levels of CysK, MetK and MetY were significantly altered in a *mcbR* knockout strain, although the levels of proteins encoded in the *cys*-operon were not affected. Thus, McbR seems to play a minor role in NCgl2640-mediated regulation.

The second hypothesis is based on the alternative annotation as a GTPase-activating protein published by Kyowa Hakko Kogyo Co. Ltd (GenBank accession no. BAC00130) and listed in ERGO, a commercial genome database that currently covers more than 600 genomes. GTPase-activating proteins are elements of GTPase-dependent regulation cascades that indirectly control gene expression (Voncken et al. 1995; Donovan et al. 2002; Litvak and Selinger 2003), mostly in eukaryotes (Caldon and March 2003) but also in bacteria (Lerner and Inouye 1991; Zhang and Inouye 2002). Thus, NCgl2640 may act as the methionine-sensing element that specifically enhances the GTP-hydrolyzing activity of its putative target GTPase that is involved in the regulation of methionine biosynthesis.

Acknowledgements We thank L. Eggeling and G. Leblon for kindly providing strain *C. glutamicum* ATCC14752 and plasmid pCGL0040, respectively.

References

Ankri S, Reyes O, Leblon G (1996a) Electrotransformation of highly DNA-restrictive corynebacteria with synthetic DNA. Plasmid 35:62–66

Ankri S, Serebrijski I, Reyes O, Leblon G (1996b) Mutations in the *Corynebacterium glutamicum* proline biosynthetic pathway: a natural bypass of the *proA* step. J Bacteriol 178:4412–4419

Bradford M (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein–dye binding. Anal Biochem 72:248–254

Caldon CE, March PE (2003) Function of the universally conserved bacterial GTPases. Curr Opin Microbiol 6:135–139

Donovan S, Shannon KM, Bollag G (2002) GTPase activating proteins: critical regulators of intracellular signaling. Biochim Biophys Acta 1602:23–45

Ebbighausen H, Weil B, Krämer R (1989) Isoleucine excretion in *Corynebacterium glutamicum*: evidence for a specific efflux carrier system. Appl Microbiol Biotechnol 31:184–190

Follettie MT, Peoples OP, Agoropoulou C, Sinskey AJ (1993) Gene structure and expression of the *Corynebacterium flavum* N13 *ask-asd* operon. J Bacteriol 175:4096–4103

Gabrielsen OS, Huet J (1993) Magnetic DNA affinity purification of yeast transcription factor. Methods Enzymol 218:508–525

Hanahan D (1983) Studies on transformation of *Escherichia coli* with plasmids. J Mol Biol 166:557–580

Hermann T (2003) Industrial production of amino acids by coryneform bacteria. J Biotechnol 104:155–172

Hermann T et al (2001) Proteome analysis of *Corynebacterium glutamicum*. Electrophoresis 22:1712–1723

Hwang BJ, Yeom HJ, Kim Y, Lee HS (2002) *Corynebacterium glutamicum* utilizes both transsulfuration and direct sulfhydrylation pathways for methionine biosynthesis. J Bacteriol 184:1277–1286

Kase H, Nakayama K (1975) L-Methionine production by methionine analog-resistant mutants of *Corynebacterium glutamicum*. Agric Biol Chem 39:153–160

Keilhauer C, Eggeling L, Sahm H (1993) Isoleucine synthesis in *Corynebacterium glutamicum*: molecular analysis of the *ilvB-ilvN-ilvC* operon. J Bacteriol 175:5595–5603

Lawrence DA, Smith DA, Rowbury RJ (1968) Regulation of methionine synthesis in *Salmonella typhimurium*: mutants resistant to inhibition by analogues of methionine. Genetics 58:473–492

Lee HS, Hwang BJ (2003) Methionine biosynthesis and its regulation in *Corynebacterium glutamicum*: parallel pathways of transsulfuration and direct sulfhydrylation. Appl Microbiol Biotechnol 62:459–467

Lerner CG, Inouye M (1991) Pleiotropic changes resulting from depletion of Era, an essential GTP-binding protein in *Escherichia coli*. Mol Microbiol 5:951–957

Leuchtenberger W (1996) Amino acids—technical production and use. In: Roehr M (ed) Biotechnology, 2nd edn. VCH, Weinheim, pp 466–504

Liebl W, Bayerl A, Schein B, Stillner U, Schleifer KH (1989) High efficiency electroporation of intact *Corynebacterium glutamicum* cells. FEMS Microbiol Lett 65:299–304

Litvak Y, Selinger Z (2003) Bacterial mimics of eukaryotic GTPase-activating proteins (GAPs). Trends Biochem Sci 28:628–631

MacNeil DJ et al (1992) Complex organization of the *Streptomyces avermitilis* genes encoding the avermectin polyketide synthase. Gene 115:119–125

Molnar-Perl I (2001) Derivatization and chromatographic behavior of the *o*-phthaldialdehyde amino acid derivatives obtained with various SH-group-containing additives. J Chromatogr 913:283–302

Mondal S, Chatterjee SP (1994) Enhancement of methionine production by methionine analogue ethionine resistant mutants of *Brevibacterium heali*. Acta Biotechnol 14:199–204

Neidhardt FC, Ingraham JL, Schaechter M (1990) Physiology of the bacterial cell: a molecular approach. Sinauer Associates, Sunderland, Mass.

Neuhoff V, Arold N, Taube D, Ehrhardt W (1988) Improved staining of proteins in polyacrylamide gels including isoelectric focusing gels with clear background at nanogram sensitivity using Coomassie brilliant blue G-250 and R-250. Electrophoresis 9:255–262

Nikiforova V, et al (2002) Engineering of cysteine and methionine biosynthesis in potato. Amino Acids 22:259–278

Rey D, Ruckert C, Kalinowski J, Puhler A, Bathe B, Hutmacher K, Pfefferle W (2002) Nucleotide sequences which code for the metD gene. In: World Intellectual Property Organization. http://www.wipo.int/ipdl/en/search/pct/search-adv.jsp

Rey DA, Puhler A, Kalinowski J (2003) The putative transcriptional repressor McbR, member of the TetR-family, is involved in the regulation of the metabolic network directing the synthesis of sulfur containing amino acids in *Corynebacterium glutamicum*. J Biotechnol 103:51–65

Rosenberg IM (1996) Protein analysis and purification. Benchtop techniques. Birkhäuser, Boston

Ruckert C, Puhler A, Kalinowski J (2003) Genome-wide analysis of the L-methionine biosynthetic pathway in *Corynebacterium glutamicum* by targeted gene deletion and homologous complementation. J Biotechnol 104:213–228

Sahm H, Eggeling L, Eikmanns BJ, Kramer R (1996) Construction of L-lysine-, L-threonine-, or L-leucine-overproducing strains of *Corynebacterium glutamicum*. Ann NY Acad Sci 782:25–39

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular cloning: a laboratory manual, 2nd edn. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.

Sauer U (2001) Evolutionary engineering of industrially important microbial phenotypes. Adv Biochem Eng Biotechnol 73:129–169

Schaefer A, Tauch A, Jaeger W, Kalinowski J, Thierbach G, Pühler A (1994) Small mobilizable multi-purpose cloning vectors derived from the *Escherichia coli* plasmids pK18 and pK19: selection of defined deletions in the chromosome of *Corynebacterium glutamicum*. Gene 145:69–73

Schägger H, von Jagow G (1987) Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Anal Biochem 166:368–379

Simic P, Sahm H, Eggeling L (2001) L-Threonine export: use of peptides to identify a new translocator from *Corynebacterium glutamicum*. J Bacteriol 183:5317–5324

Tani Y, Lim W-J, Yang H-C (1988) Isolation of L-methionine enriched mutant of a methylotrophic yeast, *Candida boidinii* No. 2201. J Ferment Bioeng 66:153–158

Voncken JW, et al (1995) Increased neutrophil respiratory burst in *bcr*-null mutants. Cell 80:719–728

Zhang J, Inouye M (2002) MazG, a nucleoside triphosphate pyrophosphohydrolase, interacts with Era, an essential GTPase in *Escherichia coli*. J Bacteriol 184:5323–5329

nature biotechnology

REVIEW

Exploiting biological complexity for strain improvement through systems biology

Gregory Stephanopoulos, Hal Alper & Joel Moxley

Cellular complexity makes it difficult to build a complete understanding of cellular function but also offers innumerable possibilities for modifying the cellular machinery to achieve a specific purpose. The exploitation of cellular complexity for strain improvement has been a challenging goal for applied biological research because it requires the coordinated understanding of multiple cellular processes. It is therefore pursued most efficiently in the framework of systems biology. Progress in strain improvement will depend not only on advances in technologies for high-throughput measurements but, more importantly, on the development of theoretical methods that increase the information content of these measurements and, as such, facilitate the elucidation of mechanisms and the identification of genetic targets for modification.

Although the term 'systems biology' entered the popular lexicon only recently, the concept of an integrated, systemic approach to the analysis and optimization of cellular processes has been applied routinely by engineers and scientists for many years. The expanded view of the cell, made possible by genome sequencing and parallel, high-throughput technologies for measuring the relative abundance of important classes of intracellular molecules, revealed the obvious: hundreds or thousands of molecules, previously excluded from the focus of research, were found to vary significantly in response to a simple genetic or environmental perturbation. Systems biology soon emerged as a term and a field of scientific inquiry to describe an approach that considers genome-scale and cell-wide measurements in elucidating biological processes and mechanisms.

Traditionally, cells were considered as elegant systems of immense complexity that were, nevertheless, well-coordinated and optimized for a particular purpose. Research efforts by necessity were narrowly focused, leading over the years to the rigorous understanding of various specific, low-level processes. Recently, genome sequencing and related technologies have, for several organisms, provided a window to the broad biomolecular landscape underlying cellular phenotype. In addition to being able to quantify the abundance of important classes of biological molecules, we can now probe the interactions among them. Systems biology aims to interpret and contextualize large, diverse sets of biological measurements and elucidate the mechanisms behind complex phenomena through an integrated perspective. This is a formidable task. To maximize the probability of success, we must anchor systems biology analyses to specific questions and build upon the existing core infrastructure created by targeted studies. Strain improvement represents a specific goal suitable for the application of systems biology. However, it remains to be seen whether the understanding gained will expedite forward engineering, that is, intelligent modifications based upon system understanding of cellular processes and justify the investment required for system characterization.

Cellular complexity is a manifestation of the enormous diversity of molecules and reaction processes needed to carry out cellular functions. This review focuses on the exploitation of complexity for strain improvement within a systems biology framework. We consider several questions of central importance to this goal, such as identifying genetic targets for modification to improve strains, elucidating biomolecular interaction networks, increasing the information content of large-scale metabolomic measurements, and integrating genomic/metabolic data. We also offer a vision for the future of systems biology in strain improvement and close with a word of caution regarding the expectations of this promising field, which has been overplayed on occasion without accounting for the many difficulties that still exist and need to be resolved before the potential of this field is fully realized.

An iterative framework for accumulating systems insight

A fundamental premise of systems-based research is that the underlying mechanisms and interactions of a biological system can be probed by introducing a variety of perturbations and measuring the system response. Figure 1 illustrates the accumulation of insight from a systems-biology cycle for strain improvement. We iteratively accumulate systems insight in two steps.

First, perturbations can be introduced into cellular networks and environments to create an altered phenotype. In recognition of the importance of these changes in probing the genotype-phenotype relationship, a diverse set of tools has emerged to create gene deletions and amplifications, which are used in conjunction with altered environments. Molecular biology advances have made it possible to perform these modifications at will. In addition to gene-specific techniques, several combinatorial tools have been developed, which, when combined with high-throughput screening, allow for randomized gene-expression levels (including deletions) and genomic library complementation[1,2].

Department of Chemical Engineering, Massachusetts Institute of Technology, Room 56-469, Cambridge, MA 02139, USA. Correspondence should be addressed to G.S. (gregstep@mit.edu).

Published online 6 October 2004; doi:10.1038/nbt1016

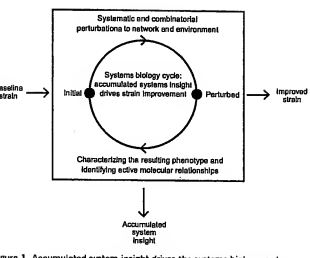


Figure 1 Accumulated system-insight drives the systems biology cycle for strain improvement. Iterative perturbations and systematic phenotype characterizations yield system insight through the integration of large data sets. A trade-off exists: more detailed characterization yields higher systems insight, but limits the number of perturbations that can be realistically evaluated.

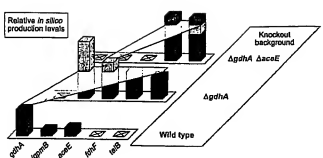


Figure 2 Identification of gene targets using global, stoichiometric modeling (H.A., Yong-Su Jin, J.M., G.S., unpublished data). This computational search makes use of a stoichiometrically balanced, genome-wide bioreaction network of *E. coli* metabolism whose fluxes are computed to maximize cell growth yield in the framework of Flux Balance Analysis and adjusted by MOMA. These results indicate that novel gene targets arise as the genotype is altered as a result of gene knockouts. This is especially evident in the case of *talB*. Although *talB* increases the production level in a *gdhA*/*aceE* knockout background, it is detrimental in a *gdhA*-only knockout background. The gene *gdhA* encodes glutamate dehydrogenase, *gpmB* encodes phosphoglucomutase, *aceE* encodes pyruvate dehydrogenase, *fdhF* encodes pyruvate formate lyase and *talB* encodes transaldolase.

Second, characterizing of the resulting phenotype allows identification of specific, differentially active molecular relationships. Using high-throughput methods, we may compare the initial and perturbed biomolecular landscapes on the gene, protein and small-molecule level. Such comparisons can reveal hundreds or thousands of molecules that vary significantly. High-level integration of biomolecular networks and states may identify concrete molecular relationships active under genetic and environmental perturbations. By specific, low-level exploration of the active biomolecular relationships, new perturbations to introduce can be identified[3].

Advances in high-throughput assays greatly enhance our ability to characterize the cellular phenotype. Sophisticated methods exist not only for analyzing and quantifying metabolite profiles, proteins and gene expression, but also for cataloging the interactions among network components. Applying these tools to a microbial system provides a detailed snapshot of cellular function.

However, one quickly encounters a trade-off between increased system characterization and the number of perturbations that can be realistically studied. Typically, the iterative strain-improvement process turns over much more rapidly than a fundamental systems biology investigation. At one extreme, one sees directed evolution in which randomly mutated strains with desired properties are preferentially selected and further modified with minimal evaluation. Here, the cellular characterization and system-wide understanding comprise only a single measurement of desired phenotype. On the other hand, even for those industrial organisms with sequenced genomes, the direct return on investment for any detailed and broad state characterization, such as gene expression profiling, can be difficult to ascertain.

Strain improvement and systems biology

Systems biology is a useful adjunct to traditional industrial programs aiming to design microbes optimized for maximal product formation. In addition to generating robust production strains, these applications often involve an important component of reverse engineering, whereby microbes with attractive properties are dissected for the purpose of transferring insights learned from their functions to the further improvement and optimization of production strains.

Global models for identification of gene targets. Advances in molecular biology for introducing genetic modifications at the molecular level have not been matched by equally effective approaches for identifying specific gene targets whose modification would bring about a desired phenotype, such as product overproduction. As a result, most strain improvement programs have resorted primarily to *ad hoc* or random approaches, whereby genetic perturbations are introduced and their effect on the phenotype of interest is evaluated along with other aspects of the resulting physiology.

The main obstacle to the rational solution of this problem is the lack of a reliable, global, metabolic model that captures the majority of the stoichiometric, kinetic and regulatory effects on metabolite interconversions and metabolic flux distributions through the cellular bioreaction network. Given the availability of genome sequences, the model closest to the above ideal is a global network of metabolic reactions described by detailed flux balances for each metabolite pool. This is usually an underdetermined system, and an *a priori* calculation of metabolic fluxes requires the optimization of a certain objective function, for example, growth yield maximization[4]. Although such 'maximum growth fluxes' do not necessarily represent the state of metabolism, they nevertheless provide insights into the distribution of carbon and energy resources required for optimal growth. A variation of the above approach, minimization of metabolic adjustment[5] (MOMA), attempts to determine more realistic flux distributions from genomic data alone by calculating profiles that are intermediate between the wild-type optimal and gene-knockout-mutant optimal.

The above framework can be used to guide the choice of gene knockout targets. In one such application of the flux balance method (modified by MOMA), we scanned the entire *Escherichia coli* genome for single- and multiple-gene knockouts that would increase the yield of lycopene (H.A., Yong-Su Jin, J.M., G.S., unpublished data). Figure 2 illustrates the selection strategy of targets for single, double and triple knockout mutants following a sequential approach. This

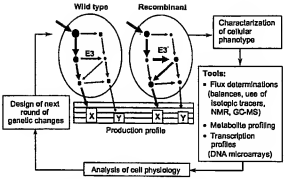


Figure 3 Exploiting complexities in metabolic networks. Metabolic engineering attempts to exploit the complexities of metabolic networks to improve cellular properties. In this schematic, wild-type cells are engineered to overexpress enzyme E3 with the goal of increasing the low yield of product Y. However, because of network interactions, overexpression of E3 has a minimal effect on the accumulation rates of either products Y or X. To improve the yield of product Y, multiple steps in the network will have to be targeted and genetically modified. In performing these steps, we gain insight into the biological system. Purple circles indicate the pool size of metabolites in the network. Arrow thickness depicts relative flux magnitude of the corresponding reactions.

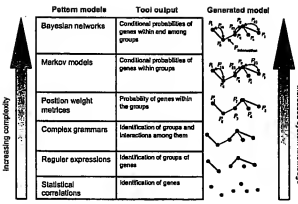


Figure 4 Approaches for pattern modeling useful in association analysis. Various pattern models and data analysis techniques may be used for the rigorous linking of datasets. In the example presented in the text, statistical correlations were used to link microarray data with phenotype (product formation). Models of increased complexity require more data, but can yield a much higher level understanding of mechanisms and underlying interactions. Furthermore, these techniques may be used to cross-validate proposed interaction networks and generate quantitative metrics (such as probabilities) for influential genes and their interactions.

method yielded a triple knockout that produced ~40% more lycopene compared with an engineered overproducing *E. coli* strain.

The success of the above approach justifies some optimism regarding more rational approaches to strain improvement but also raises a host of interesting questions. First, can we expect all gene knockouts identified by the above methods to show increased productivity? The answer is clearly no. The above computational approach simply identifies mutants with an increased availability of precursors for product formation, but as a purely stoichiometric calculation, it cannot incorporate regulatory or kinetic effects. Second, will such sequential, steepest ascent–like methods always cover the entire landscape? We clearly do not know. In this particular case, all possible gene pairs were investigated in an exhaustive computational search, and no pair that had not been covered by the sequential approach was found. This is a non-generalizable result, nevertheless. These observations underscore the importance of the iterative scheme of Figure 1 in elucidating the complex landscape of cellular function.

Randomized genetic tools, such as overexpression libraries and transposon mutagenesis, also allow broad-range perturbations of cellular systems. When coupled with targets identified from global modeling, these methods can be used for effective strain improvement. Furthermore, characterizing these systems allows the dissection of critical subnetworks within the cell. By investigating how product formation correlates with these regulatory networks, putative molecular interactions may be inferred and tested in subsequent perturbations.

Bioreaction network analysis. Metabolic engineering aims to improve strains using modern genetic tools. We showed that strains can be modified by introducing specific transport, conversion or deregulation changes that result in flux redistribution and product yield improvement[6]. Metabolic engineering differs from genetic engineering in that it is concerned with the entire metabolic system rather than with the overexpression of a gene. Although genetic engineering can successfully overexpress or otherwise modify a gene, this may have little impact on cell physiology. By examining the properties of the metabolic network in its entirety, metabolic engineering attempts to identify targets for amplification as well as to assess rationally the effect of such changes on the properties of the network. Metabolic engineering is thus a progenitor of functional genomics and systems biology in that it represents the first organized effort to reconstruct and modify pathways using genomic tools and guided by information about the behavior of the entire system[7]. This approach is illustrated by the concept of the distribution of kinetic control (metabolic control analysis, MCA)[8,8] and by the modification of multiple enzymes to achieve flux amplification, as we recently verified experimentally[9].

Figure 3 shows how the general paradigm of Figure 1 is applied to the specific case of metabolic engineering. A key point here is that the first rounds of genetic changes, though rarely successful, can nevertheless contribute invaluable insights that can guide the strain improvement cycle to a successful conclusion. Figure 3 also lists the tools available for more elaborate evaluation of strain physiology after a perturbation, although how these tools can be used systematically for strain improvement is still unclear. Nevertheless, we have shown that flux determination is an essential component of strain evaluation for metabolic engineering[10]. The examples below proceed from the modification of simple, linear product pathways to the perturbation of complex interacting networks encountered in central carbon metabolism.

We have investigated the production of threonine in a lysine-producing strain[11]. Carbon flux from aspartate is distributed between the lysine and the threonine-isoleucine pathways at the aspartate semialdehyde (ASA) branch point. There are three reactions from ASA to threonine, and the product of the first one, homoserine dehydrogenase (HDH), is feedback-inhibited by threonine. The identification of a mutant gene encoding a threonine-insensitive HDH solved the feedback-inhibition problem, but its overexpression succeeded in redirecting only 50% of the aspartate flux towards threonine. The reason was an imbalance between the activities of HDH and the next enzyme in the pathway, homoserine kinase (HK). When these two enzymes were balanced in a new construct that incorporated inducible control of the HK expression, almost 100% of the flux was redirected toward threonine synthesis. In this case, enzyme activity measurements along with enzymatic kinetic data about the sequential pathway reactions were the additional information that guided subsequent research.

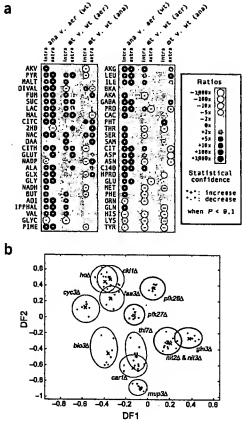


Figure 5 Use of metabolic state assays for phenotypic characterization. (a) GC-MS measurements of a broad diversity of intracellular metabolites 3B (J.M. and S. Villas-Boas, unpublished data). Here, we observe a ratio for each metabolite for three environmental and genetic perturbations (environmental: aerobic and aerobic cultivation, genetic: wild-type yeast and *gdh1* knockout), intracellularly and extracellulary. Statistically significant differences in metabolite pools are indicated by a '+' or '–' to indicate increases or decreases, respectively. By comparing these metabolite levels across samples, we may infer activated genetic and metabolic pathways. For instance, the anaerobic cultivations accumulate much higher levels of TCA cycle intermediates because the corresponding TCA cycle reactions occur at slower rates. (b) GCMS peaks for metabolite data may be clustered to classify mutants[37]. Here, 19 yeast deletion strains with various knockouts affecting central carbon metabolism and amino acid production were cultivated overnight in microtiter plates. From GC-MS data for the extracellular metabolites, a degrees of analysis (DFA) model was trained with 20 principal components. Based upon the first two principal components, the plot demonstrates clusters of samples that correspond to certain mutants In other words, the metabolic footprint of gas chromatography-mass spectrometry peaks contains enough phenotypic information to differentiate among mutants.

In a larger network application, modification of central carbon metabolism was considered for overproduction of aspartic-acid family amino acids, such as lysine. Single-gene overexpression had only marginally increased production[12], but flux control coefficient calculations revealed that most of the kinetic control for lysine overproduction resided in the lysine pathway[12]. Additionally, as shown by our laboratory, isotopic tracer probes applied to well-designed genetic backgrounds identified pyruvate carboxylase as the key reaction supplying more than 90% of the lysine carbon[13]. This led to the sequencing of pyruvate carboxylase in *Corynebacterium glutamicum*; however, its overexpression only marginally increased the yield of lysine. Further investigation identified aspartokinase as the bottleneck enzyme in a pyruvate carboxylase–overexpressing strain: when aspartokinase was simultaneously overexpressed with pyruvate carboxylase, the specific lysine productivity rose by one- to threefold depending on the carbon source used. In this example, detailed flux calculations provided the critical information that was best used in an MCA-defined framework.

The above case studies represent the different types of input information sought to explain the observed strain physiology and to rationally guide strain improvement. There are numerous other applications of this rapidly expanding field, whose very definition is linked with systems biology and which has produced impressive results in the past decade[14–29].

Integrating physiological and transcriptional data. The rapid adoption of new technologies suggests that future developmental programs will have the benefit of large volumes of data for the characterization of extensive libraries of strains. A key question is how to use these data most profitably to identify gene targets for modification. A related issue of critical importance is how to link data from the expression and metabolic phenotypes to help reveal the genetic basis of strain performance and organism physiology in general.

Various mathematical constructs are available to create links between large data sets and phenotypes. In one such application, association discovery was used to evaluate a library of unsequenced fungal strains of *Aspergillus terreus* for their ability to overproduce the antibiotic lovastatin[30]. First, using gene overexpression, a large number of strains with diverse lovastatin and (+)-Geodin production profiles was generated, and the strains were characterized by metabolite and transcriptional profiling. From this wealth of biological data, Askenazi *et al.*[30] extracted key putative parameters and genes by a statistical association analysis. To do this, each transcriptional data ratio and metabolic profile ratio (both calculated using the parental strain as basis) was correlated with the product level using Pearson correlations and principal components analysis (PCA). These statistical tools, applied to the sign of the correlation coefficient, provided a measure of the relative strength and impact (either positive or negative) that each transcript had on the product level. The analysis revealed key trends for some genes associated either positively or negatively with the production of lovastatin and (+)-Geodin. This led to the identification of genes whose modulation alters lovastatin production and eventually to an improvement in production of over 50%.

Techniques for linking data sets are not limited to simple statistical metrics, as illustrated in the above example. Various pattern discovery and characterization tools can yield different outcomes and levels of knowledge about the systems. Figure 4 summarizes some commonly used pattern models and the typical system detail extracted from such analyses. Given enough data, one can extract probabilistic models that completely capture cellular interactions. In this way, relevant biomolecular networks may be discovered and manipulated without prior knowledge of an interaction network, which is especially important for unsequenced or poorly studied microorganisms. These concepts are illustrated in Box 1.

Use of metabolites for phenotype characterization. Strain phenotype characterization has relied primarily on transcript abundance and protein measurements. Only rarely have small metabolites been included in the measurement set, reflecting the difficulties in sampling and analyzing these molecules owing to their rapid time scales of change, especially at the high-throughput rates envisioned by the systems biology cycle of Figure 1. However, it is now accepted that, by

Box 1 Identification of active biomolecular pathways

Researchers continue to expand and refine pathway databases, but how does one determine which of the molecular relationships that define a mechanism might be relevant under a particular set of conditions? In one sense, pathway databases reflect a road map of possible functional routes connecting cellular components—but it is a static, noisy, incomplete road map that by itself yields limited understanding of how a perturbation induces a particular cellular phenotype. Clearly, we want to dissect how pathways are activated or repressed when a perturbation is introduced to identify targets for genetic manipulation.

Entries in pathway databases vary considerably in quality and reflect observations from either specific, bottom-up studies or global, top-down interaction screens. For instance, years of studying the galactose genetic switch has yielded a wealth of literature and very accurate interaction data about core regulation. Other connections, however, might be less well validated. *E. coli* pathway databases are based exclusively on knowledge accumulated from years of bottom-up studies. For yeast, however, the knowledge from bottom-up studies is supplemented by global experimental data. For protein-protein binding data, yeast two-hybrid[31] and protein-complex mass spectrometry experimentation have provided multifold coverage of each pairwise combination of protein interactions. Likewise, for transcription factor–mediated gene regulation, chromatin immunoprecipitation[32] microarrays measure DNA promoter binding sites genome-wide. These top-down approaches reveal interactions—albeit for a limited collection of organisms because of high costs—that have not been the focus of a specific, bottom-up study.

The determination of active biomolecular interactions can be facilitated with the Cytoscape software platform, which allows the integration of transcriptional data through global, biomolecular-interaction databases to identify the most active pathways[33]. Ideker *et al.*[34] demonstrated that differential (transcriptional) state measurements placed in the context of biomolecular networks can generate valuable pathway insight. Using galactose-pathway perturbations, they showed that putatively interacting genes were more likely to be differentially active together. This result reinforces the concept that pathways are activated and deactivated in a coordinated fashion. The Cytoscape platform allows visualization of these differential state data on the biomolecular network. Often, the user inputs differential gene expression data and uses the curated Biomolecular Interaction Network Database (BIND) for protein-protein (usually signaling) and protein-DNA (usually regulatory) interactions. Going a step further, the ActiveModules plug-in searches the biomolecular networks to determine which pathways are most affected by a set of perturbations[35]. By processing global data, one finds potentially interesting pathways that may be further explored through lower-level, detailed modeling and experiments.

Thus, a Cytoscape analysis facilitates the completion of the systems biology cycle by allowing the identification of active molecular relationships to target with the next iteration's set of perturbations. Figure 6 illustrates a specific example in which differential transcription data reveal the pathways activated and deactivated upon the removal of the GAL80 repressor protein.

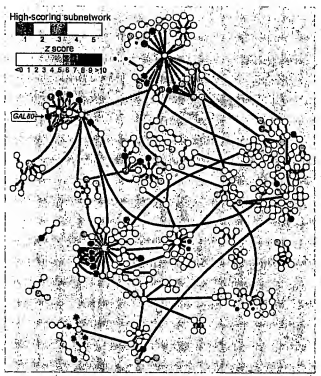


Figure 6 Determining active pathways after removing a transcription factor repressor. Identification of active pathways helps define gene targets. In this experiment, the GAL4 repressor gene GAL80 was deleted. Using microarray data superimposed on a predetermined set of protein-protein (pp) and protein-DNA (pd) interactions for yeast, the differential gene expression after gene deletion reveals the corresponding activated subnetwork illustrated above. Even without galactose present, removal of the GAL80 triggers cellular galactose-processing pathways by eliminating the repression of the GAL4 transcription factor. Node color indicates differential expression statistical significance for the particular gene, whereas node outline color and interaction edges between nodes indicates activated subnetworks. Significance of differential expression does not distinguish between upregulation and downregulation states; thus, both GAL80 (here, eliminated) and GAL1 (here, upregulated after GAL80 removal) will possess high z-scores as differentially active. As is evident from this figure, a single modulation of a gene can have a cascade effect throughout the biomolecular interaction network.

The ActiveModules plug-in scores each gene node and then searches, scores and compares possible active subnetworks. We note that, although the primary active pathway includes *GAL80* and the surrounding genes, not all active pathways directly connect to *GAL80*. This observation highlights the noisy, incomplete nature of the biomolecular networks. Often, one cannot directly trace the cascading effects of a perturbation because only a subset of biomolecules (e.g., genes but not small molecules) and their interactions have been considered. However, by identifying relationships that become active after certain perturbations, one gains valuable insight regarding rational gene targets to disrupt at the next iteration.

globally assaying metabolic states, we can identify a more diverse set of active molecular relationships indicative of mechanisms that encompass stoichiometry *and regulation*, as small metabolites are critical in regulating higher-level (transcriptional, translational) processes. Unfortunately, small molecules possess a wider range of chemical characteristics than do transcripts and are more difficult to measure simultaneously.

Gas chromatography coupled to mass spectrometry allows high-throughput analysis at relatively low cost. The gas chromatograph separates metabolites, whereas the mass spectrometer identifies and quantifies the metabolites corresponding to a given peak. Metabolite profiles in the culture supernatant are used to define a footprint of metabolic processes that occurred intracellularly. Such metabolite profiling will be applied increasingly to better define the cellular phenotype.

How will the observed metabolic fingerprints or footprints profiles help elucidate the physiological state of organisms following some perturbation? Figure 5 illustrates the use of dimensional reduction methods for visualizing different physiological states in a lower-dimensional space. Fisher discriminant analysis[36] projection of the metabolic data collected for strains subjected to environmental (aerobic and anaerobic cultivation) and genetic (*gdh1* knockout) perturbations succeeds in classifying the samples collected under different conditions and from different strains. Supervised and unsupervised classification methods such as these will be important in defining the location of a desirable metabolic state (such as one of high productivity) and in developing bioreactor controls that lead the system to the desired state in the course of a process.

Metabolic data can also be incorporated into databases that integrate transcription, protein-protein or protein-DNA interactions, and metabolism to identify biomolecular subnetworks that become activated in response to a perturbation. This vision entails an expanded Cytoscape framework encompassing metabolism along with transcriptional and higher-level processes in the cellular hierarchy. This broader metabolic state characterization will allow better understanding of the interplay between different pathways and will enhance the confidence of mechanism identification through the use of an expanded and more diverse set of measurements. As such, these methods have a role in the reverse engineering of strains.

Words of caution

The plethora of new data and new analytic methods justifiably leads to ambitious goals for systems biology. After many years of focused reductionist research, a deluge of information is expanding the scope of most investigations and promises to transform glimpses into snapshots of a dynamic world where cells grow, divide and produce, or organisms develop, differentiate and begin to deviate from the norm. No one can deny the opportunities that present themselves, but one must also be mindful that the problem that we set out to address is several orders of magnitude larger than those with which we are familiar. Consequently, it is important that we temper our expectations of immediate results and not lose sight of the following points:

First, despite the wealth of available genomic data, we are still unaware of numerous genes involved in important interactions and processes. A common misconception is that the genomic effort and accompanying analysis are almost complete. However, as recent results have demonstrated, this could not be further from the truth: genomic maps are continuously updated by discarding or adding (occasionally substantial amounts of) genes, proteins and new biomolecular interactions for pathways that were considered well understood.

Second, constructing biomolecular networks demands significant resources and expertise. Biomolecular networks incorporate a multitude of relationships connecting several types of components. At the genome scale, interaction maps require large experimental investments and subsequent analysis and curation. For instance, global protein-protein interaction maps exist for only a handful of model species, and reconstructing well-studied and well-documented networks, such as metabolic pathways, in a genome context requires years of curation.

Third, completeness is an elusive goal even for the better understood biomolecular networks. As more genomes are sequenced, the effort to uncover the structure and function of genetic regulatory networks has given birth to many databases, each of which attempts to distill the most salient features from incomplete and at times flawed knowledge. We mentioned the scant agreement among individual yeast two-hybrid screens. Clearly, 'accepted' interactions vary significantly across databases and over time, even for yeast, the best-studied system. Furthermore, many databases do not distinguish among direct and indirect interactions, raising uncertainty about the design of experiments to disrupt such interactions.

Fourth, high-throughput analytical tools, such as DNA microarrays, are widely available only for conducting measurements at the transcriptional level, and even these require training and incur significant costs. Going further to measure protein levels, protein states, regulatory elements, metabolites and metabolic fluxes requires complex and specialized equipment and significant expertise. Consequently, partnerships and collaborations are a *sine qua non* for effective research in systems biology.

Finally, patience is advised, as the more complex hypotheses derived from systems approaches are disproportionately harder to validate. Typically, after a perturbation, differential gene expression and network searching reveal active putative biomolecular networks that span dozens, if not hundreds, of genes. Figure 6 shows that even for a small network, a single gene removal causes a cascade of activation and deactivation among many genes. Although we can demonstrate cause and eventual effect, verifying the specific mechanism of each gene's activation/deactivation becomes a large task.

Looking forward

The reorientation of the research frame of mind to accept systems approaches to biological research will lead to better technologies for probing cellular phenotypes at ever increasing levels of accuracy and resolution. One can safely assume that the present trend of high-throughput measurement development will continue unabated, addressing molecules that are more difficult to measure and incorporating exciting new micro- and nanotechnologies.

More and more measurements, however, will not ensure faster progress toward the above objectives in the absence of effective tools for increasing the information content of measurements. Without underestimating the challenges in the development of measurement technologies, we believe that the key limiting step will be the development of mathematical-computational approaches that organize and integrate data to answer specific biological and biotechnological questions. This article has reviewed methods for meeting this fundamental challenge. It is our hope that it will help engage systems-oriented researchers and thus spur more activity in this area.

ACKNOWLEDGMENTS

We acknowledge the DuPont-MIT Alliance for research support and Kyle Jensen for his helpful suggestions.

COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/naturebiotechnology/

1. Hemmi, H. *et al.* Identification of genes affecting lycopene formation in *Escherichia coli* transformed with carotenoid biosynthetic genes: candidates for early genes in isoprenoid biosynthesis. *J. Biochem.* **123**, 1088–1096 (1998).
2. Badarinarayana, V. *et al.* Selection analyses of insertional mutants using subgenic-resolution arrays. *Nat. Biotechnol.* **19**, 1060–1065 (2001).
3. Ideker, T. & Lauffenburger, D. Building with a scaffold: emerging strategies for high to low-level cellular modeling. *Trends Biotechnol.* **21**, 255–262 (2003).
4. Edwards, J.S. & Palsson, B. The *Escherichia coli* MG1655 In silico metabolic genotype: Its definition, characteristics, and capabilities. *Proc. Natl. Acad. Sci. USA* **97**, 5528–5533 (2000).
5. Segre, D., Vitkup, D. & Church, G.M. Analysis of optimality in natural and perturbed metabolic networks. *Proc. Natl. Acad. Sci. USA* **99**, 15112–15117 (2002).
6. Stephanopoulos, G., Aristidou, A. & Nielsen, J. *Metabolic Engineering: Principles and Methodologies* (Academic Press, San Diego, 1998).
7. Stephanopoulos, G. & Vallino, J.J. Network rigidity and metabolic engineering in metabolite overproduction. *Science* **252**, 1675–1681 (1991).
8. Kacser, H. & Burns, J.A. The control of flux. *Symp. Soc. Exp. Biol.* **27**, 65–104 (1973).
9. Kotfas, M.A., Jung, G.Y. & Stephanopoulos, G. Engineering metabolism and product formation in *Corynebacterium glutamicum* by coordinated gene overexpression. *Metab. Eng.* **5**, 32–41 (2003).
10. Stephanopoulos, G. Metabolic Fluxes and Metabolic Engineering. *Metab. Eng.* **1**, 1–10 (1999).
11. Colon, G.E., Nguyen, T.T., Jetten, M.S., Sinskey, A.J. & Stephanopoulos, G. Production of isoleucine by overexpression of *ilvA* in a *Corynebacterium lactofermentum* threonine producer. *Appl. Microbiol. Biotechnol.* **43**, 482–488 (1995).
12. Cremer, J. *et al.* Control of the lysine biosynthesis sequence in *corynebacterium glutamicum* as analyzed by overexpression of the individual corresponding genes. *Appl. Environ. Microbiol.* **57**, 1746–1752 (1991).
13. Kotfas, M.A., Jung, G.Y., Aon, J.C. & Stephanopoulos, G. Effect of pyruvate carboxylase overexpression on the physiology of *Corynebacterium glutamicum*. *Appl. Environ. Microbiol.* **68**, 5422–5428 (2002).
14. Berrios-Rivera, S.J., Bennett, G.N. & San, K.-Y. The effect of increasing NADH availability on the redistribution of metabolic fluxes in *Escherichia coli* chemostat cultures. *Metab. Eng.* **4**, 230–237 (2002).
15. Cameron, D.C. *et al.* Metabolic engineering of propanediol pathways. *Biotechnol. Prog.* **14**, 116–125 (1998).
16. Causey, T.B., Shanmugam, K.T., Yomano, L.P. & Ingram, L.O. Engineering *Escherichia coli* for efficient conversion of glucose to pyruvate. *Proc. Natl. Acad. Sci. USA* **101**, 2235–2240 (2003).
17. Farmer, W.R. & Liao, J.C. Improving lycopene production in *Escherichia coli* by engineering metabolic control. *Nat. Biotechnol.* **18**, 533–537 (2000).
18. Fussenegger, M. & Betenbaugh, M.J. Metabolic engineering II. Eukaryotic systems. *Biotechnol. Bioeng.* **79**, 509–531 (2002).
19. Kaup, B. *et al.* Metabolic engineering of *Escherichia coli*: construction of an efficient biocatalyst for D-mannitol formation in a whole-cell biotransformation. *Appl. Microbiol. Biotechnol.* **64**, 333–339 (2004).
20. Khosla, C. & Bailey, J.E. Heterologous expression of a bacterial hemoglobin improves the growth properties of recombinant *Escherichia coli*. *Nature* **331**, 633–635 (1988).
21. Lee, S.Y. & Lee, Y. Metabolic engineering of *Escherichia coli* for production of enantiomerically pure (*R*)-(-)-hydroxycarboxylic acids. *Appl. Environ. Microbiol.* **69**, 3421–3426 (2003).
22. Martin, V.J.J. *et al.* Engineering the mevalonate pathway in *Escherichia coli* for production of terpenoids. *Nat. Biotechnol.* **21**, 796–802 (2003).
23. Nakamura, C. & Whited, G. Metabolic engineering for the microbial production of 1,3-propanediol. *Curr. Opin. Biotechnol.* **14**, 454–459 (2003).
24. Ostergaard, S. *et al.* Increasing galactose consumption by *Saccharomyces cerevisiae* through metabolic engineering of the GAL gene regulatory network. *Nat. Biotechnol.* **18**, 1283–1286 (2000).
25. Snell, K. & Peoples, O. Polyhydroxyalkanoate polymers and their production in transgenic plants. *Metab. Eng.* **4**, 29–40 (2002).
26. Stafford, D.E. *et al.* Optimizing bioconversion pathways through systems analysis and metabolic engineering. *Proc. Natl. Acad. Sci. USA* **99**, 1801–1806 (2002).
27. Stephanopoulos, G. & Kelleher, J. Biochemistry: how to make a superior cell. *Science* **292**, 2024–2025 (2001).
28. Viswanathan, K. *et al.* Engineering sialic acid synthetic ability into insect cells: Identifying metabolic bottlenecks and devising strategies to overcome them. *Biochemistry* **42**, 15215–15225 (2003).
29. Xue, Y. & Sherman, D.H. Alternative modular polyketide synthase expression controls macrolactone structure. *Nature* **403**, 571–575 (2000).
30. Askenazi, M. *et al.* Integrating transcriptional and metabolite profiles to direct the engineering of lovastatin-producing fungal strains. *Nat. Biotechnol.* **21**, 150–156 (2003).
31. von Mering, C. *et al.* Comparative assessment of large-scale data sets of protein-protein interactions. *Nature* **417**, 399–403 (2002).
32. Lee, T.I. *et al.* Transcriptional regulatory networks in *Saccharomyces cerevisiae*. *Science* **298**, 799–804 (2002).
33. Shannon, P. *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).
34. Ideker, T. *et al.* Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* **292**, 929–934 (2001).
35. Ideker, T., Ozier, O., Schwikowski, B. & Siegel, A.F. Discovering regulatory and signaling circuits in molecular interaction networks. *Bioinformatics* **18** (suppl. 1) 233–240 (2002).
36. Stephanopoulos, G., Hwang, D., Schmitt, W.A., Misra, J. & Stephanopoulos, G. Mapping physiological states from microarray expression measurements. *Bioinformatics* **18**, 1054–1063 (2002).
37. Allen, J. *et al.* High-throughput classification of yeast mutants for functional genomics using metabolic footprinting. *Nat. Biotechnol.* **21**, 692–696 (2003).

Biotechnology Letters (2005) 27: 1135–1140
DOI 10.1007/s10529-005-8463-y

A high-throughput method for screening of rapamycin-producing strains of *Streptomyces hygroscopicus* by cultivation in 96-well microtiter plates

Zhi-nan Xu*, Wen-he Shen, Xi-yang Chen, Jian-ping Lin & Pei-lin Cen
Institute of Bioengineering, Department of Chemical Engineering and Bioengineering, Zhejiang University, 310027, Hangzhou, China
**Author for correspondence (Fax: +86-571-87951220; E-mail: znxu@zju.edu.cn)*

Received 18 March 2005; Revisions requested 5 April 2005; Revisions received 25 May 2005; Accepted 30 May 2005

Key words: bioassay, high-throughput method, rapamycin, *Streptomyces hygroscopicus*, 96-well microtiter plate

Abstract

A novel high-throughput cultivation method was developed to rapidly screen large numbers of rapamycin-producing mutants of *Streptomyces hygroscopicus* by duplicate culturing of isolates on the surfaces of agar-solidified 96 wells in microtiter plates. One copy of the cultures was used for the rapamycin bioassay and the other identical copy, representing potentially high yielding strains, was preserved for further study. By integrating 96-well solid cultivation and the bioassay, we screened more than 7000 isolates and found 10 high-yielding strains. From these, one mutant produced 420 μg rapamycin/ml, which was double the yield of parent strain used in the submerged fermentation process.

Introduction

Rapamycin is an antifungal antibiotic produced by *Streptomyces hygroscopicus* (Sehgal *et al.* 1975, Vezina *et al.* 1975). It also has potent immunosuppressive activity by inhibiting T-cell activation and proliferation (Biere *et al.* 1990, Dumont *et al.* 1990). It has been approved in United States for use in combination with cyclosporine and corticosteroids for various treatments (Kahan *et al.* 1991, Kaplan *et al.* 1998). Compared with its competitive agents, the productivity of rapamycin in *Streptomyces hygroscopicus* is very low, being only about 200–300 mg/l (Kojima *et al.* 1995, Lee *et al.* 1997, Chen *et al.* 1999). Further improvement in productivity of this industrial microorganism is imperative for commercial success.

For the past decade, new technologies have been developed to improve productivity of industrial strain, such as combinatorial biosynthesis and genome shuffling (Zhang *et al.* 2002, Donadio *et al.* 2003, Petri *et al.* 2004). However, traditional mutation is still used to improve industrial strains. A search for a high-yielding strain depends largely on the total number of mutants that can be screened after mutagenesis (Vinci *et al.* 1991). Even with genome shuffling strategy, positive mutants must be screened out from a large number of isolates obtained by using recursive protoplast fusion (Zhang *et al.* 2002). A variety of screening methods have been developed targeting different types of microorganisms (Filtenborg *et al.* 1983, Smedsgaard *et al.* 1997, Bragulat *et al.* 2001). For antibiotic-producing strains, the "agar plug" method was widely used in 1980s, and many modified versions were developed for preliminary screening with high efficiency (Du Toit *et al.* 2000, Kumar *et al.* 2000).

The aim of this work was to develop a high-throughput screening (HTS) technique to isolate the high rapamycin-producing strains from a large number of mutants after nitrosoguanidine treatment of *Streptomyces hygroscopicus*. The combination of 96-well microtiter-plates cultivation

technique and agar-plug method was used to make the screening procedure more efficient. The productivity by rapamycin by the isolated strains was then examined by submerged fermentation.

Materials and methods

Microorganisms and mutation

Streptomyces hygroscopicus ZX-S–21 was isolated after mutation of *S. hygroscopicus* ATCC 29253 and preserved in our laboratory. This low-level rapamycin producer was grown at 28 °C for 20 days on solid medium (oat meal 20 g/l, agar 20 g/l, pH 7.0). Spores were harvested in 15 ml sterile 0.85% saline containing 0.1% Tween 80 and 2 ml of this suspension was mixed with 0.2 ml of 5 mg nitrosoguanidine (NTG)/ml, and killing ratio of ca. 90% was obtained by shaking the mixture for 1 h at 150 rpm. The spore suspension, containing 10–1000 c.f.u., was prepared. For each Petri dish, 0.2 ml of the spore suspension was spread onto the surface of SYPC medium (soluble starch 10 g/l, yeast extract 6 g/l, peptone 6 g/l, Casamino acids 1.5 g/l, $MgSO_4$ 0.5 g/l, K_2HPO_4 1.0 g/l, pH 6.3–6.8). After cultivation for 10 days at 28 °C, separate colonies appeared with a large number of spores.

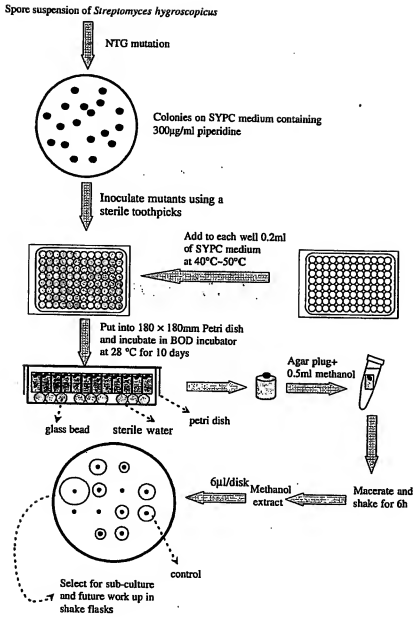
Candida albicans (ATCC 11651) was used in the bioassay of rapamycin. It was grown at 28 °C for 2 days on an agar slant (glucose 20 g/l, yeast extract 10 g/l, peptone 20 g/l, agar 20 g/l, pH7.0).

Cultivation of isolates on 96-well agar array and extract of rapamycin

Each well in aseptic 96-well microtiter plates was filled with 0.2 ml sterile SYPC medium containing 1.5% agar. After solidification, the spores from one colony in the Petri dish were transferred into one corresponding well of the duplicate microtiter plates (A and B). After every well was inoculated, the 96-well plates were put into a 180 × 180 mm Petri dish in which 100 ml sterile water and 25 pieces of glass beads were added to prevent the agar column from drying. The dishes were incubated at 28 °C for 10 days in a biochemical oxygen demand (BOD) incubator. Then, each medium column in one 96-well plate (A) was plugged out totally and transferred into a 1.5 ml Eppendorf tube in which 0.5 ml methanol was pre-added. After 6 h shaking at 30 °C, methanol extracts were centrifuged for 15 min at 5000 g, and the supernatant was then used for bioassay. Another 96-well plate (B) was stored in refrigerator at 4–6 °C. If the strain in one well of plate (A) showed high rapamycin productivity by bioassay, the strain in the corresponding well of plate (B) will be further validated by submerged fermentation. Control plugs were also prepared with *S. hygroscopicus ZX-S–21* and harvested using the same procedure.

Bioassay method for rapamycin

Rapamycin production was assayed using a paper disc-agar diffusion method using *Candida albicans* ATCC 11651 as the indicator (Kojima *et al.* 1995, Lee *et al.* 1997). The agar medium dispensed in a 90 × 90 mm Petri dish composed of two separate layers. First, 15 ml test medium (glucose 5 g/l, peptone 2 g/l, agar 20 g/l) formed a base layer in the Petri dish then, after solidification, 0.1 ml suspension of *C. albicans* mixed with 5 ml of sterile saline containing 0.8% agar at 40–50 °C was immediately poured onto the base layer to constitute the upper one. Then 12 pieces of 6 mm diam. paper disk were placed onto each dish, and 6 μl methanol extract from the agar plug of plate (A) was transferred onto 10 paper disks respectively. In the remaining two paper disks, one containing 6 μl pure methanol was used as negative standard, the another containing 6 μl methanol extract from the parent strain was used as positive standard. In order to facilitate rapamycin in the methanol extract diffusing into the medium of Petri dish, the plates were stored at 4 °C for 12 h, then incubated for 2 days at 30 °C. The inhibitory zone was observed and the diameter was measured. The larger the diameter of inhibitory zone is, the higher the concentration of rapamycin in methanol extract will be; therefore, the strain with the largest inhibitory zone can be easily picked out and the corresponding strain in plate (B) will be further examined by shake-flask submerged fermentation. The detailed procedure of this high-throughput screening technique is illustrated in Figure 1.



C. albicans bioassay plate showing zones of inhibition

Fig. 1. An integrated process with a modified agar plug method and *Candida albicans* bioassay.

Shake flask screening for high-yielding isolates

Spores in plate (B) picked out according to bioassay were inoculated into a 250 ml Erlenmeyer flask containing 30 ml seed medium (soluble starch 10 g/l, yeast extract 6 g/l, peptone 6 g/l, Casamino acids 1.5 g/l, $MgSO_4$ 0.5 g/l, K_2HPO_4 1 g/l, pH 7.0). Incubation was conducted at

28 °C for 60 h on a rotary shaker (220 rpm), and 3 ml of the resulting culture broth was transferred into a 250 ml Erlenmeyer flask containing 30 ml medium. Fermentation was continued for 5 days at 28 °C on a rotary shaker (240 rpm). Then 2 ml of fermentation broth was centrifuged for 15 min at 4000 *g*. The cell pellet was mixed with the same volume of methanol and shaken intensively for 4 h. After centrifugation at 4000 *g* for 15 min, the supernatant was subjected to HPLC analysis to determine rapamycin concentration. An Agilent 1100 series HPLC system equipped with a Hypersil BDS C_{18} reversed-phase analytical column (250 × 4.6 mm I.D., 5 μm particle) was used. The mobile phase consisted of methanol/water (78:22, v/v) at 1 ml/min and the eluate was monitored at 277 nm. Before analysis, the samples needed to be diluted properly, and rapamycin in the sample was estimated using an internal standard.

Results and discussion

Determination of cultivation time of isolates in 96-well microtiter plate

The 96-well microtiter plate is able to culture many mutants on the surface of agar-solidified medium simultaneously with uniform conditions. The mutant of *S. hygroscopicus* grew well on the plate for more than 20 days. In order to determine the optimal cultivation duration for rapamycin production, the spores of *S. hygroscopicus ZX-S–21* were used to inoculate three 96-well microtiter plates in parallel. At intervals, three samples were taken out from 96-well plates and extracted with methanol. The productivity of rapamycin was determined by the diameter of inhibitory zones against *C. albicans* ATCC 11651. The results were shown in Figure 2. The discs with pure methanol, used as a negative control, showed no inhibitory zone. The inhibitory zones with *S. hygroscopicus ZX-S–21* were round with sharp edge. After cultivation of *S. hygroscopicus* for four days, rapamycin started to be synthesized and the inhibitory zone size was expanded quickly with time and stopped to increase until the 10th day. Further extending of culture time did not increase the inhibitory zone

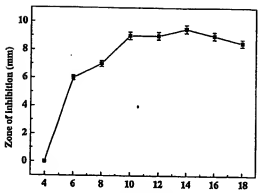


Fig. 2. The relationship between the diameter of inhibition zone and the incubation time of *S. hygroscopicus* on the surface of 96 solidified agar wells. All the experiments were carried out in triplicate, and values are the average of three independent determinations.

size and even decreased after 14th day. Based on above observations, in the following experiments, the cultivation time in the 96-well microtiter plate was fixed for 10 days.

Solvent extraction of rapamycin from the agar plug

Rapamycin is an intracellular product. To examine its inhibitory zone, it is necessary to extract rapamycin from mycelium of *S. hygroscopicus* grown on the agar plugs in the 96-well microtiter plate. Various organic solvents were evaluated and methanol showed the highest efficiency. The time course of methanol extraction was shown in Figure 3. Complete extraction of rapamycin needed 6–8 h. Efficient extraction of rapamycin by methanol from *S. hygroscopicus* pellets takes 2 h (Lee *et al.* 1997). The extension of extraction time is probably caused by the mass transfer resistance when the mycelia are partially wrapped by the agar. In the following experiments, methanol extraction for 6 h was adopted for further bioassay.

Development of HTS method for bioassay of rapamycin

Although *C. albicans* ATCC 11651, used as test organism, is sensitive to rapamycin, the sensitivity

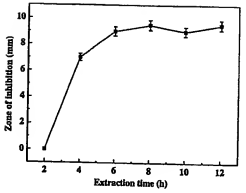


Fig. 3. The relationship between the extract time of rapamycin from an agar plug and the diameter of inhibition zone. All the experiments were carried out in triplicate, and values are the average of three independent determinations.

and consistency of inhibitory zone size is dependent on the growth period. After detailed experimental evaluation, the best results were obtained that *C. albicans* ATCC 11651 in double-layer medium should be incubated for 16–18 h at 28 °C.

After above optimization processes of key operation conditions, one high throughput screening (HTS) procedure with 96-well formats was successfully developed in our laboratory, and was used to screen high rapamycin-producing isolates.

A total of 2108 separate colonies appeared in SYPC medium after treatment of spore suspension of *S. hygroscopicus ZX-S–21* with NTG. Of these, 2008 colonies grew well in 96-well microtiter plates; those with poor or no growth were discarded. By running the HTS procedure, those isolates providing the largest inhibition zones over the original one were selected as a population of positive. The size of inhibitory zones from selected isolates were measured and recorded. The respective duplicate agar plugs (B) were then subcultured to validate rapamycin productivity in shake-flask fermentation. Each isolate was incubated in triple and the rapamycin productivity in each flask was analyzed by HPLC. The final rapamycin concentration was the average of that in three parallel flasks. The results showed that three isolates (N102, N437, N758), which gave the largest sizes of inhibitory zones, were the top-three rapamycin producers in the shake-flask fermentation.

The spores of these three mutants were mixed together and used for next-run mutation and screening. After three-run similar work, 10 high-yielding strains were screened out from more than 7000 mutants. There was an almost linear relationship between inhibitory zone sizes of the 10 high-yielding strains and the corresponding rapamycin production in the shake-flask culture (Figure 4). Strain N5632 was the most productive one from three rounds of the HTS procedure and the rapamycin yield reached 420 mg/l, which is the highest value ever reported in the literature in submerged fermentation and is double that produced by the original ZX-S–21. A 124% improvement of rapamycin productivity was achieved by traditional mutation breeding, however, the highest rapamycin productivity of mutant C14–2 was only 139 mg/l (Cheng *et al.* 2001).

The advantages of the HTS method developed in this work are its high throughput and ease of isolating high producing strains. Within 3 months, three rounds of HTS can be performed and more than 7000 isolates being evaluated. Also expensive and tedious HPLC analysis can be reduced to the minimum. The new HTS procedure is versatile for the screening of high producing strains for antibiotic production and may be applied in the other screening systems after improvement. In this laboratory, the HTS procedure is being applied in the genome shuffling.

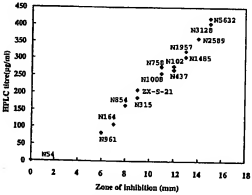


Fig. 4. Zones of inhibition of *C. albicans* against HPLC titre of rapamycin in shake flask. All the experiments were carried out in triplicate, and values are the average of three independent determinations.

References

Bierer BE, Mattila PS, Standaert RF, Herzeoberg LA, Burakoff SJ, Crabtree G, Schreiber SL (1990) Two distinct signal transmission pathways in T lymphocytes are inhibited by complexes formed between an immunophilin and either FK506 or rapamycin. *Proc. Natl. Acad. Sci. USA.* 87: 9231–9235.

Bragulat MR, Abarca ML, Cabanes FJ (2001) An easy screenioq method for fungi producing ochratoxin A io pure culture. *Int. J. Food Microbiol.* 71: 139–144.

Chen YL, Krol J, Sterkin V, Fan WM, Yan XK, Huang WM, Cioo JL, Julien C (1999) New process control strategy used in a rapamycin fermentation. *Proc. Biochem.* 34: 383–389.

Cheng YR, Huaog J, Qiang H, Lin WL, Demain AL (2001) Mutagenesis of the rapamycin producer *Streptomyces hygroscopicus* FC904. *J. Antibiot.* 54: 967–972.

Donadio S, Sosio M (2003) Strategies for combioatorial biosynthesis with modular polyketide synthases. *Comb. Chem. High Throughput Screen* 6: 489–500.

Du Toit EA, Rauteobach M (2000) A sensitive standardized micro-gel well diffusion assay for the determination of antimicrobial activity. *J. Microbiol. Methods* 42: 159–165.

Dumoot F, Melino J, Staruch MJ, Koprak SL, Fischer PA., Sigal NH (1990) Distinct mechanisms of suppression of murine T-cells activation by the related macrolides FK506 and rapamycin. *J. Immunol.* 144: 251–258.

Filtenborg O, Frisvad JC, Svendsen JA (1983) Simple screen-method for molds producing intracellular mycotoxins io pure cultures. *Appl. Environ. Microbiol.* 45: 581–585.

Kahan BD, Gibbons S, Tejpal N, Stepkowski SM, Chou TC (1991) Synergistic interactions of cyclosporioe and rapamycin to iohibit immune performances of normal human peripheral blood lymphocytes in vitro. *Transplantation* 51: 232–239.

Kaplan B, Meier-Krieache HU, Napoli KL, Kahan BD (1998) The effects of relative timioq of sirolimus and cyclosporine microemulsion formulation coadministration on the pharmacokinetics of each agent. *Clin. Pharmacol. Ther.* 63: 48–53.

Kojima I, Cheog YR, Demain AL (1995) Carbon source nutritioo of rapamycio biosynthesis in *Streptomyces hygroscopicus*. *J. Ind. Microbiol. Biotechnol.* 14: 436–439.

Kumar M.S, Kumar PM, Sarnaik HM, Sadhukhan AK (2000) A rapid technique for screening of lovastatio-producing strains of *Aspergillus terreus* by agar plug and Neurospora crassa bioassay. *J Microbiol. Methods* 40: 99–104.

Lee MS, Kojima I, Demaio AL (1997) Effect of nitrogeo source on biosynthesis of rapamycin by *Streptomyces hygroscopicus*. *J. Ind. Microbiol. Biotechnol.* 19: 83–86.

Petri R, Schmidt-Dannert C (2004) Dealing with complexity: evolutionary engineering and genome shuffling. *Curr. Opin. Biotechnol.* 15: 298–304.

Sehgal SN, Baker H, Vezina C (1975) Rapamycin (AY–22,989), a new antifungal aotibioiic: II. Fermentatioo, isolatioo, characterization. *J. Antibiot.* 28: 726–727.

Smedsgaard J (1997) Micro-scale extractioo procedure for standardized screeniog of fungal metabolite productioo io cultures. *J Chromatogr. A* 760: 264–270.

Vinci VA, Hoerner TD, Coffmann AD, Schimmel TG, Dabora RL, Kirpeker AC, Ruby CL, Stieber RW (1991) Mutants of a lovastatin hyper producing *Aspergillus terreus* deficient in the production of sulochrin. *J. Ind. Microbiol.* 8: 113–120.

Vezina C, Kudelski A, Sehgal SN (1975) Rapamycio (AY–22,989), a new antifungal antibiotic: I. Taxonomy of the producing Streptomycete and isolation of the active principle. *J. Antibiot.* 28: 721–726.

Zhang YX, Perry K, Vinci VA, Powell K, Stemmer WP, Cardayre SB (2002) Genome shuffling leads to rapid phenotypic improvement io bacteria. *Nature* 415: 644–646.

Solid-phase high-throughput screening of enzyme variants: Detecting enhanced nitrilase activity

*William J. Coleman, Steven J. Robles, and Mary M. Yang**

Kairos Scientific Inc.
10225 Barnes Canyon Road #A110
San Diego CA 92121
USA
Phone (858) 626-8170
Fax (858) 626-8177
Web www.kairos-scientific.com

**Corresponding author: myang@kairos-scientific.com*

Introduction

The use of enzymes to synthesize pharmaceutical intermediates has been the focus of considerable research and development effort over the past decade, as companies seek to produce novel chemical structures while reducing the manufacturing cost and eliminating chemical waste. Enzymes and whole cells are attractive catalysts not only because of their regioselectivity and enantioselectivity, but also because they can perform biocatalytic reactions under mild conditions with fewer synthesis steps[1]. However, the stability, selectivity, and productivity of wild-type enzymes are not always optimal for chemical processing.

Development of enzyme-based processing has been hindered due to low activity on nonnatural substrates, enzyme instability, and poor or inappropriate enantioselectivity. Molecular biology tools now exist to remedy these deficiencies, including the application of directed evolution to engineer novel enzymes with enhanced properties[2–5].

Directed evolution generally consists of the following sequence of steps: (1) mutating the gene encoding the enzyme to create a large population of variants, (2) expressing the gene products in a host organism, (3) screening for the desired properties, and (4) retrieving the desired variants. These steps can be repeated until the ideal enzyme variant is produced.

Identifying enzyme variants with improved characteristics requires instrumentation and methods that provide greater throughput, information content, and flexibility to efficiently process highly complex gene libraries. Using nitrilase as an example, we describe a newly developed technology that can be used to perform these tasks.

Nitrilases and pharmaceutical synthesis

A number of key drugs or their pharmaceutical precursors contain functional groups consisting of carboxylic acids connected to chiral centers. These include, for example, the cholesterol-lowering drug Lipitor® (atorvastatin) and the analgesic Aleve® ((S)-naproxen). Methods for converting nitrile derivatives to yield chiral products using conventional chemistry have the disadvantage of requiring harsh chemicals that generate undesirable waste products. Consequently, enzyme-based synthetic routes have been developed using nitrilases to convert nitrile-containing precursors into the (R)- and (S)-enantiomers of the corresponding carboxylic acid[1,6,7] *(Figure 1)*.

Up to now, high-throughput screening of nitrilase libraries was typically performed in a liquid-phase format using standard 96- or 384-well microplates. By utilizing colored indicators or fluorogenic reagents, the reaction products can be detected by plate readers equipped with absorbance or fluorescence detectors[6]. However, microplate-based systems require robotic handling and dispensing to operate in a high-throughput mode. This type of operation requires large numbers of plates and large volumes of potentially difficult-to-synthesize or expensive substrates—which nevertheless may only facilitate screening rates of several thousand variants per day. Increasing the screening rate to 10^5–10^6 variants per instrument per day while simultaneously simplifying the sample handling and minimizing reagent usage will make it possible to screen much larger mutagenized-enzyme libraries.

Figure 1. Nitrilase-catalyzed hydrolysis of nitriles into carboxylic acids

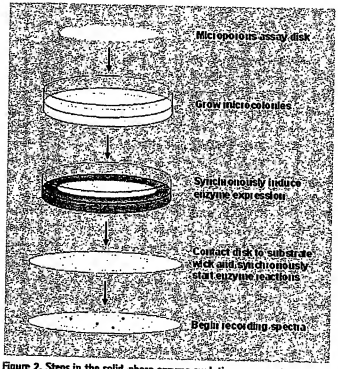


Figure 2. Steps in the solid-phase enzyme evolution assay system

Kcat technology for solid-phase enzyme screening

To achieve these enhancements, Kairos Scientific (San Diego, CA) has developed an integrated system that uses a high-density, solid-phase format in combination with digital imaging spectroscopy (DIS) to screen enzyme variants expressed in microcolonies. DIS combines image processing and optical spectroscopy so that complete spectral information can be obtained for every pixel or feature in a target image. The commercial platform for enzyme screening is known as Kcat™ Technology and includes assays, instrumentation, and computer algorithms. This technology[2,8] has been employed to engineer the properties of industrial enzymes[3,9]. Kcat Technology can be used to monitor enzymatic activity on tens of thousands of different enzyme variants simultaneously. This is accomplished by acquiring full spectral and/or kinetic information from microcolonies that are simultaneously undergoing a color-forming reaction catalyzed by the enzymes they express.

The basic solid-phase assay employs a colony-forming microbial host to express the gene(s) of interest. *Figure 2* shows an example of how the assay is performed using plasmids expressed in *E. coli*. Briefly:

1) *E. coli* cells transformed with plasmids containing the mutagenized library are randomly deposited on a circular, microporous membrane (the assay disk).

2) The assay disk is transferred to LB agar containing antibiotics, and the cells are incubated at 37°C overnight until they form microcolonies.

3) The assay disk is transferred to agar medium containing an inducer (e.g., IPTG) to synchronously induce expression of the enzyme in all the cells. (This step is optional.)

4) The assay disk is transferred to a wick containing the substrate to synchronously initiate the enzyme reaction. This is done inside the temperature-controlled Kcat device (e.g., at 37°C) to begin the measurement.

5) Images are acquired over time at the appropriate wavelength(s) to monitor the progress of the enzymatic activity simultaneously in all of the microcolonies on the assay disk. After the data is analyzed using the Kcat software, microcolonies displaying the desired activity are picked, and their corresponding plasmids are retrieved and purified.

As another demonstration of this technology, a library of random nitrilase variants was generated by error-prone PCR mutagenesis and expressed in *E. coli* microcolonies. The solid-phase assay was performed on the substrate at 37°C to detect those mutants with the highest activity based on the increase in absorbance over time (due to ammonia production). Images were acquired every 60 seconds.

Figure 3 shows a composite image of the assay disk with the reacted microcolonies and the Kcat Graphical User Interface, or GUI. The assay disk is 47 mm in diameter and contains approximately 9,000 individual microcolonies. The GUI has three interactive windows that can be used to visualize and analyze the assay data, which in this case consists of approximately 15,000 kinetic traces. The kinetics for each pixel are sorted (*Figure 4*) and displayed as thin horizontal lines in the Contour Plot Window (*Figure 3*, right). Absorbance of each pixel at a given time is color-coded from black/blue (low) to pink/white (high) according to the color scale at the bottom.

Several pixels representing the highest activity were selected from this window by clicking and dragging the mouse over this top por-

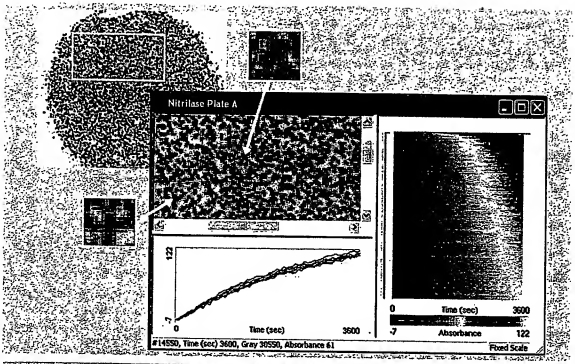


Figure 3. Screening a library of nitrilase variants using the Kcat instrument. The two most active micro-colonies identified by the software are highlighted in red. Microcolonies with zero activity are not visible in the image. Assay disk diameter is 47 mm. Approximately 9,000 individual variants are on the disk.

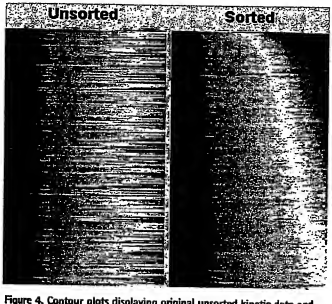


Figure 4. Contour plots displaying original unsorted kinetic data and data sorted by maximum absorbance

tion of the Contour Plot. These pixels are identified by the thin red tic mark to the left of the Contour Plot Window. The pixels selected from this window are automatically highlighted in the Image Window (top left) to identify the most active variants. The corresponding kinetics for these selected pixels are also automatically displayed in the Plot Window (bottom left).

In addition to a screening throughput of over one million variants per instrument per day, Kcat Technology provides a number of added benefits. A solid-phase format based on microcolonies has very small volume demands (ca. 50 nl). The microcolonies can be intact, lysed, or pretreated so that the enzyme is secreted. Assay design is flexible and may include coupled enzyme assays. High-molecular-weight substrates, which cannot penetrate cell walls or are not amenable to pipetting, can also be used. Due to the fact that the assay disk can be manipulated without disturbing the cells, variants can be simultaneously evolved for multiple properties by subjecting them to various physical and chemical pretreatments (heating, pH changes, etc.). The capability of acquiring full spectra over time means that the reactions can be multiplexed. This feature is convenient for comparing the enzyme activity on multiple substrates in order to change the speci-

ficity[2]. Likewise, acquisition of both spectral and time-based information enables simultaneous evolution of kinetic properties as well.

The libraries to be screened using Kcat Technology can be generated in a number of ways. Mutagenized DNA libraries can be created using error-prone PCR mutagenesis[10] to introduce random changes throughout the protein, or by a combinatorial cassette technique such as recursive ensemble mutagenesis (REM) and related methods[11-13]. The latter techniques are extremely efficient for targeting specific regions of the sequence, such as an antibody-binding site or an enzyme active site. REM and its related methods have been shown to achieve a 10-millionfold improvement in the proportion of functional and unique variants as compared to purely random combinatorial mutagenesis. Kcat screening methods can also be applied to recombinant protein libraries for optimizing protein therapeutics and to environmental or genomic libraries for discovering new genes.

Conclusion

Application of Kcat Technology to enzyme screening will make it feasible to employ much more intensive mutagenesis methods for engineering projects involving biocatalysts such as nitrilase. Simplifying the sample handling and raising the throughput will facilitate the assaying of much larger, more complex libraries to identify enzymes that are active on nonnatural substrates and that have improved stability. A user-friendly system will enable individual laboratories to create their own customized enzymes for specific applications. This in turn will facilitate the development of new synthetic processes to create enantiopure chemicals for the pharmaceutical industry.

Acknowledgement

This work was supported under NIH R44GM55470 and Kairos IR&D funds.

REFERENCES

1. Sheldon RA and van Rantwijk F. Biocatalysis for sustainable organic synthesis. *Austral J Chem* 57:281-289 (2004).

2. Bylina EJ et al. Solid-phase enzyme screening. *ASM News* 66:211-217 (2000).

3. Delagrave S et al. Application of a very high-throughput digital imaging screen to evolve the enzyme galactose oxidase. *Protein Eng* 14:261-267 (2001).

4. Powell KA et al. Directed evolution and biocatalysis. *Angew Chem Int Ed Engl* 40:3948-3959 (2001).

5. DeSantis G et al. Creation of a productive, highly enantioselective nitrilase through gene site saturation mutagenesis (GSSM). *J Am Chem Soc* 125:11476-11477 (2003).

6. Banerjee A, Kaul P, Sharma R, and Banerjee UC. A high-throughput amenable colorimetric assay for enantioselective screening of nitrilase-producing microorganisms using pH sensitive indicators. *J Biomol Screen* 8:559-65 (2003).

7. Robertson DE et al. Exploring nitrilase sequence space for enantioselective catalysis. *Appl Environ Microbiol* 70:2429-2436 (2004).

8. Coleman WJ, Robles SJ, and Yang MM. Engineering enzymes at high throughput. *Gen Eng News* 23:34, 39 (2003).

9. Delagrave S et al. Combinatorial mutagenesis algorithms, digital imaging spectroscopy, and solid-phase assays for directed evolution. In: *Enzyme Functionality*, Svendsen A (ed.), pp. 507-523, Marcel Dekker, New York (2004).

10. Leung DW, Chen E, and Goeddel DV. A method for random mutagenesis of a defined DNA segment using a modified polymerase chain reaction. *Technique* 1:11-15 (1989).

11. Goldman ER and Youvan DC. An algorithmically optimized combinatorial library screened by digital imaging spectroscopy. *Bio/Technology (NY)* 10:1557-1561 (1992).

12. Delagrave S and Youvan DC. Searching sequence space to engineer proteins: Exponential ensemble mutagenesis. *Bio/Technology (NY)* 11:1548-1552 (1993).

13. Delagrave S, Goldman ER, and Youvan DC. Recursive ensemble mutagenesis. *Protein Eng* 6:327-331 (1993).